

# **Un aplicativo en Excel para la enseñanza y modelación de una serie de tiempo:**

## **Precio externo del café colombiano**

**Oscar Mauricio Monsalve Tamayo**



**UNIVERSIDAD TECNOLÓGICA DE PEREIRA**  
**Facultad ciencias básicas**  
**Maestría en la enseñanza de la matemática**  
**Pereira**  
**Marzo 16 del 2016**

## **AGRADECIMIENTOS:**

Quiero agradecer a Dios, primero que todo, quien me ha dado salud y la oportunidad de terminar este trabajo de tesis de maestría, a mis padres: María Nilsa Tamayo y Oscar Monsalve Castaño a quienes debo todo mi ser y la motivación de profundizar mis conocimientos en matemáticas, a mi esposa Gloria Gómez, al profesor Ricardo López Varona Msc en física, quien fue mi profesor de matemáticas **V** y laboratorio de Física **II** en el pregrado, en la Universidad Tecnológica de Pereira, quien además me tendió la mano cuando no tenía asesor de tesis y con sus sabios consejos pude terminar este trabajo de investigación, a la doctora Gladys Elena Salcedo Echeverry por sus aportes a mi trabajo, También a mi amigo Oscar zapata.

## TABLA DE CONTENIDO

	Pág.
1. Marco teórico	
Resumen	9
1.1 Introducción	10
1.2 Justificación	11
1.3 Planteamiento del problema	12
1.4 Objetivos generales y específicos	14
1.5 Hipótesis	15
1.6 Metodología	16
2. Conceptos básicos de las series de tiempo	17
2.1 Proceso estocástico	17
2.2 Proceso estacionario	19
2.3 Proceso de ruido blanco	19
2.4 Función de Autocorrelación	19
2.5 Función de Autocorrelación Parcial	20
3. Modelos	21
3.1 Formulación general de modelos lineales	21
3.2 Pasos en la construcción de modelos Arima	22
Paso 1: Identificación de los términos del modelo	22
Paso 2: Estimación de los parámetros del modelo	22
Paso 3: Validación del modelo	23
Paso 4: Predicción	24
3.3. Formulación general de Modelos no lineales	25

3.3.1 Modelo Arch	25
3.3.1.2 Modelo Arch (1)	26
3.3.1.3 Modelo Arch ( r)	28
3.3.2 Modelo Garch	29
3.3.2.1 Modelo Garch (1,1)	29
3.3.2.2 Modelo Igarh (1,1)	30
3.3.2.3 Modelo Egarch (1,1)	31
3.3.2.4 Modelo Tarch (1,1)	30
3.4 Construcción de los modelos no lineales	33
Paso 1: Identificación de los términos del modelo	32
Paso 2: Estimación de parámetros del modelo	33
Paso 3: Diagnóstico	34
4. Estudio de la serie de tiempo precio externo del café colombiano	36
4.1 Análisis exploratorio de la serie precio externo del café colombiano	36
4.2 Estudio de la volatilidad de la serie precio externo del café colombiano	
De aplicarle logaritmo y una diferencia.	40
4.3 Prueba de Heterocedasticidad.	46
5. Modelación de la serie precio externo del café colombiano en Eviews 3.1	47
5.1 Selección de modelo	47
5.2 Pronósticos	51
5.3 Comparativo de los resultados de los pronósticos del aplicativo en	
Excel con los arrojados por Eviews 3.1.	52
5.4 Modelación en Matlab.	52

Conclusiones	61
Bibliografía	62
Anexo: Manual de usuario Aplicativo en Excel	63

Índice de gráficas	Pág.
• Figura 4.1: Histograma de la serie Z	35
• Figura 4.1.1 : Gráfico Quantile-Quantile	36
• Figura 4.1.2 : Gráfico de la Desviación Stándard de la serie café	36
• Figura 4.1.3 :Correlograma precio externo del café	37
• Figura 4.1.4 :Gráfica del precio del café y la serie de retornos	39
• Figura 4.2.1 :Correlograma de retornos	42
• Figura 4.2.2 : Histograma de la serie Dlog(z)	43
• Figura 5.4.1 : Parámetros Modelo Arch 1 Matlab	51
• Figura 5.4.2 : Parámetros Modelo Garch (1,1)	52
• Figura 5.4.3 : Parámetros Modelo Tarch (1,1)	53
• Figura 5.4.4 : Parámetros Modelo Egarch (1,1)	54
• Figura 5.4.5 : Gráfica del precio externo del café en Matlab	55
• Figura 5.4.6: Gráfica de la serie después aplicar logaritmos y una diferencia en Matlab.	55
• Figura 5.4.7 : Test de Dickey Fuller	56
• Figura 5.4.8 : Test de efectos Arch	56
• Figura 5.4.9: Gráfica de Autocorrelación FAC de la serie precio externo del café colombiano después de haber aplicado logaritmo y una diferencia.	57

- Figura 5.5: Autocorrelación Parcial de la serie precio externo del café colombiano. 57
- Figura 5.6 : FAC y FACP de residuos del Modelo Tarch (1,1) 58
- Figura 5.7: Función de Autocorrelación simple y parcial del Modelo Tarch (1,1) 58

## Índice de Tablas

	Pág.
• Tabla 4.1 : Modelo AR( 1)	38
• Tabla 4.2 : Test de Dickey Fuller para la serie $Z_t$	38
• Tabla 4.3 : Test de Dickey Fuller para la serie de los retornos	43
• Tabla 4.4 : Regresión de la serie contra la constante	44
• Tabla 4.5 : Prueba de Heterocedasticidad de la serie $Z_t$	
• Tabla 5.1 : Ajuste Modelo AR(1) + ARCH(1) de la serie de los retornos	46
• Tabla 5.1.1: Ajuste del Modelo AR (1) + GARCH (1,1) a la serie de los retornos.	47
• Tabla 5.1.2: Ajuste del Modelo AR (1) + EGARCH (1,1) a la serie de los retornos.	48
• Tabla 5.1.3: Ajuste del modelo AR (1) + TARCH (1,1) a la serie de los retornos.	49
• Tabla 5.1.4 Elección del mejor Modelo	49
• Tabla 5.2 Pronósticos del Modelo TARCH (1,1)	50
• Tabla :5.3 Comparativo Pronósticos	51
• Tabla 5.4.1 : Criterio Akaike Modelo ARCH 1	52
• Tabla 5.4.2 : Criterio Akaike Modelo Garch(1,1)	52
• Tabla 5.4.3: Criterio Akaike Modelo Tarch (1,1 )	53
• Tabla 5.4.4 Criterio Akaike Modelo Egarch (1,1)	54



• Tabla 5.4.5: Resumen de los criterios Akaike para Modelos	54
• Tabla 5.4.6.8: Criterios del Modelo Tarch (1,1)	59
• Tabla 5.4.6.9 : Pronósticos Modelo Tarch (1,1)	59
• Tabla 5.4.7 : % de error del Modelo Tarch (1,1) usando Matlab	59

## Resumen

En el presente trabajo se hace un análisis detallado de la serie de tiempo: precio externo del café colombiano desde la perspectiva de los modelos Autorregresivos, usando un aplicativo en Excel.

Inicialmente se hace un recorrido de los modelo lineales ARIMA, y de los modelos no lineales Arch, Garch, Tarch, debido a que esta serie financiera presenta efectos Arch, es decir heterocedasticidad condicionada, además presenta mucha volatilidad, es asimétrica y leptocúrtica. Luego se ajustan los modelos utilizando el Eviews, el Matlab y el aplicativo realizado en Excel.

Palabras clave: GARCH, ARIMA, heterocedasticidad condicionada, volatilidad.

## 1.1 Introducción

En el presente trabajo se analiza la serie precio externo del café colombiano desde las perspectivas de las series de tiempo, destacando la importancia de los modelos con heterocedasticidad condicionada y la importancia de los pronósticos para las series financieras.

Inicialmente se presenta una revisión de los modelos ARIMA, GARCH (heterocedasticidad condicional autorregresiva generalizada) y se dan algunas propiedades de los procesos con sus demostraciones. Además, se presenta una aplicación de los modelos Arima-Garch, con el software Eviews 3.1 considerando la serie de tiempo: precio externo del café colombiano que va desde enero de 1980 hasta diciembre del 2015. También, se muestra el modelo de predicción más adecuado para estimar el precio externo y la volatilidad del café colombiano, usando el software Eviews 3.1 y Matlab, con el objetivo de comparar los pronósticos con un aplicativo en Excel el cual analiza la serie con los métodos de medias móviles, mínimos cuadrados y suavización exponencial.

En los Capítulos 2 y 3 se describen algunos modelos lineales y no lineales para series de tiempo, en el Capítulo 3 se describe estadísticamente la serie: precio externo del café colombiano, en el Capítulo 4 se encuentra el mejor modelo para la serie y se realizan los pronósticos para enero del 2016 usando el software Eviews 3.1 y el Matlab, finalmente se hace un comparativo entre las predicciones arrojadas por el modelo y los resultados del aplicativo en Excel.

## 1.2 Justificación

Las series de tiempo constituyen uno de los más importantes instrumentos teóricos y a su vez herramientas para la modelación e interpretación de fenómenos económicos registrados en el tiempo. Es por esto que debido al avance de los computadores, la metodología de los modelos Arima formalizados por Box y Jenkins en 1976, han hecho posible mediante cuatro etapas (identificación – estimación –verificación –pronóstico) llegar a modelar series de tiempo tanto a nivel de economía, astronomía, medicina, etc.

Ejemplos de series de tiempo los encontramos en:

- i. Agricultura: Cuando observamos las cosechas anuales y los precios de un producto.
- ii. Economía: Cuando se presentan intereses semanales, indicadores de precios mensuales de una acción.
- iii. Ingeniería: Cuando se observa el sonido, señales eléctricas y voltajes.
- iv. Geofísica: Turbulencia, ondas en los océanos.
- v. Medicina: Encefalogramas, electrocardiogramas.
- vi. Meteorología: Velocidad de los vientos, temperatura diaria, lluvias anuales.
- vii. Sociales: tasa de nacimiento anual, tasa de mortalidad, tasa de accidentes, tasa de homicidios.

El análisis de una serie de tiempo requiere entonces de aplicativos que puedan ayudar en la enseñanza de las series de tiempo.

### 1.3 Planteamiento del Problema

Las series de tiempo también llamadas series cronológicas o series históricas, son un conjunto de datos numéricos que se obtienen en períodos regulares y específicos a través del tiempo; los tiempos pueden darse en años, meses, semanas, días u otra unidad adecuada al problema que se esté trabajando.

Ejemplo de las series de tiempo son: ventas mensuales de un producto en una empresa, producción total o anual de petróleo durante un cierto número de años, o las temperaturas anunciadas cada hora por el servicio meteorológico para un aeropuerto.

Uno de los objetivos de las series de tiempo es hacer proyecciones o pronósticos sobre la actividad futura, suponiendo establecer las condiciones y variaciones, registrándolas hasta la fecha, la cual permite planear y tomar decisiones a corto y largo plazo.

Para el análisis de series de tiempo existen softwares de ayuda como el Eviews 3.1, Matlab, Spss, R, Splus, Minitab, etc...

Microsoft Excel es muy importante ya que es un software potente, flexible y además uno de los más usados en el mundo; no existe otro programa que pueda acercársele en cuanto a funciones o flexibilidad; su ámbito de aplicabilidad va desde la economía, a la sicología, de la biología a las ciencias sociales, de las matemáticas aplicadas, a la administración de recursos humanos. Es una herramienta muy importante, sobre todo para las empresas ya que en este software se pueden realizar la contabilidad, análisis de sus facturas, análisis de datos, agenda de clientes y todo tipo de base de datos. Además podemos crear desde tablas y/o gráficos estáticos o dinámicos, hasta macros. En el mundo miles de millones de dólares se mueven gracias a este programa. Muchas decisiones se toman apoyadas en él. Millones de empresas de todo el mundo, simplemente no podrían funcionar si no tuvieran Excel en sus equipos de cómputo. Gran parte de los programas a la medida de “independientes” que existen, en realidad utilizan esta herramienta como motor de “cálculo”. Ahora en la enseñanza podemos considerar el Excel como una herramienta poderosa, que no sólo nos permite llevar la parte contable, sino proveer magníficos ambientes para el estudio del modelado de problemas, para el uso de expresiones matemáticas y para la solución de diversos problemas en varias disciplinas.

Por otra parte, El software Eviews entre muchas operaciones y funciones estadísticas el que realiza, ofrece varias herramientas de análisis y modelación de series de tiempo, entre ella la familia de modelos no lineales con heterocedasticidad condicionada

En este trabajo de grado se utiliza el Excel para la enseñanza de las series de tiempo, ya que debido a la diversidad de funciones que posee es muy didáctico a la hora de tratar temas estadísticos.

Para enfrentar el problema difícil y complejo de modelar y enseñar series de tiempo debe hacerse uso de aplicativos en Excel y Matlab que permitan aprender a identificar modelos rápidamente y realizar los respectivos pronósticos.

## 1.4 Objetivo General

Desarrollar un aplicativo en Excel para la enseñanza y modelación de una serie de tiempo como el precio externo del café colombiano, en centavos de dólar por libra de 435.6 gramos excelso desde enero de 1980 hasta diciembre del 2015.

### Objetivos específicos

- Mostrar a través del software Excel la manera de modelar una serie de tiempo, como el precio externo del café colombiano y su importancia a nivel de facilidad, en la enseñanza de la estadística.
- Realizar un comparativo entre los resultados del pronóstico arrojados por el aplicativo en Excel y los arrojados con el Eviews 3.1 y el Matlab.
- Destacar la importancia de los modelos Arch (Autorregresive Condicional Heterocedastic) y Garch (Generalized Autorregresive Conditional Heterocedastic) para el análisis de series financieras y en especial para la serie precio externo del café colombiano.

## 1.5 Hipótesis

Suponemos que se dispone de datos en intervalos regulares de tiempo (meses) y deseamos utilizar en lo posible la “inercia” en el comportamiento de las series para proveer su evolución futura. Este tipo de análisis lo denominamos invariante, porque utiliza como única información la propia historia de la serie, basándose en que las condiciones futuras serán análogas a las pasadas. Recordemos que los modelos invariantes son especialmente útiles para la previsión a corto plazo, la previsión debe tener en cuenta la evolución de las otras variables, relacionadas con lo que nos interesa, prever mediante un modelo regresión dinámica o modelo de la función de transferencia.



## 1.6 Metodología

Se revisó un buen número de libros enfocados al tema del análisis de series de tiempo, artículos en internet y algunos manuales para el software Eviews 3.1. El contenido del trabajo está basado principalmente en el libro Econometría de Domodar N.Gujarati y Dawn C Poter, Mc Graw Hill, 5ª Ed, Capítulo XXI en adelante; sin embargo, tiene temas que requieren mayor explicación para un mejor comprensión de la materia. Por ejemplo, el texto mencionado, no discute con detalle el tema de los modelos invariantes, por lo que se debió ampliar con el libro: Introducción al análisis de series temporales de José Alberto Mauricio, Universidad complutense de Madrid.

Los conjuntos de datos que se utilizaron para ejemplarizar la teoría vienen de la federación de cafeteros, página web:

[http://www.federacióndecafeteros/particulares/.../119\\_estadísticas\\_histórica](http://www.federacióndecafeteros/particulares/.../119_estadísticas_histórica).

# CAPÍTULO 2

## CONCEPTOS BÁSICOS DE LAS SERIES DE TIEMPO

### 2.1 proceso estocástico y series de tiempo

Un proceso estocástico es un conjunto de variables aleatorias que dependen de un argumento,  $t \in \tau \subseteq R$ . En el análisis de series de tiempo el argumento  $t$  se refiere al tiempo. También lo podemos definir como una familia de variables aleatorias  $\{Y_t\}$ , donde  $t$  denota el tiempo, tales que para cualquier elección finita de valores  $t_1, t_2, \dots, t_n$  existe la distribución de probabilidad conjunta correspondiente a las variables aleatorias  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ .

Una de las formas de describir un proceso estocástico de manera sencilla es a partir de los momentos poblacionales, particularmente con los de primer y segundo orden.

i.  $E(Y_t) = \mu_t, \forall_t$

ii.  $V(Y_t) = \sigma_t^2, \forall_t$

iii.  $\gamma_{t,s} = \text{Cov}(Y_t, Y_s) = E[Y_t - \mu_t][Y_s - \mu_s], \forall_{t,s}$

Una serie de tiempo es una realización de las variables aleatorias de un proceso estocástico referidas a un conjunto índice  $T$ . En el contexto de series de tiempo, el conjunto índice es el tiempo, el cual puede ser discreto o continuo. Mientras un proceso estocástico es la colección de las variables aleatorias, una serie de tiempo es una realización finita de un proceso estocástico. Es decir, la serie de tiempo es el resultado de observar la colección de las variables aleatorias. De esta manera, existe un número infinito de realizaciones (series de tiempo) de un mismo proceso estocástico.

Una serie de tiempo presenta variaciones en sus valores observados. Dichas variaciones se deben a cuatro clases de causas a las cuales se les denomina las componentes de la serie de tiempo. Estas cuatro componentes son: componente de tendencia, componente cíclico, componente estacional y componente aleatorio, de tal forma que cualquier serie de tiempo puede descomponerse en estas componentes.

- Componente de tendencia ( $T_t$ ):

Es la componente a largo plazo que constituye la base del crecimiento o declinación de una serie histórica. Las fuerzas básicas que producen o afectan la tendencia de una serie son: cambios en la población, inflación, cambio tecnológico e incremento en la productividad.

- Componente estacional ( $E_t$ ):

Se denomina componente estacional o estacionalidad a aquellas fluctuaciones subanuales (por ejemplo mensuales, trimestrales) ya que se repiten regularmente de año en año. Las tres características más importantes del fenómeno estacional son: la repetición con cierta regularidad, es posible medirlo y separarlo de las otras componentes que influyen en el movimiento de la serie. Es causado principalmente por fuerzas no económicas, exógenas al sistema económico, que los tomadores de decisiones no pueden controlar o modificar a corto plazo. Las causas más frecuentes de estacionalidad suelen ser debidas a un factor climático.

- Componente cíclico ( $C_t$ ):

Es un conjunto de fluctuaciones en forma de ondas o ciclos, de más de un año de duración.

- Componente aleatorio ( $A_t$ ):

Son movimientos erráticos que no siguen un patrón específico y que obedecen a causas diversas. Este componente es prácticamente impredecible.

Cuando observamos una serie de tiempo en una gráfica o en una tabla, estos componentes no son visibles. Lo que se observa es la conjunción de ellos a la vez en cada punto de la serie. Generalmente se asumen tres modelos: aditivo, modelo multiplicativo, modelo logaritmo aditivo, para la representación de una series temporal en sus cuatro componentes

i. Modelo aditivo:

$Y_t = T_t + C_t + E_t + A_t$ . Este modelo asume que los componentes de la serie son independientes.

ii. Modelo multiplicativo:

$Y_t = T_t \times C_t \times E_t \times A_t$ . Este modelo asume que las componentes están interrelacionados.

iii. Modelo Log Aditivo:

$$\text{Log}(Y_t) = T_t + C_t + E_t + A_t$$

## 2.2 Proceso estacionario

Un proceso estocástico  $\{Y_t\}$  con  $E(Y_t^2) < \infty$ , es estacionario de segundo orden o en sentido amplio si satisface las siguientes tres condiciones:

- i.  $E(Y_t) = \mu_t$ , para todo  $t$ . Es decir la media del proceso  $Y_t$  es constante.
- ii.  $V(Y_t) = \sigma^2 < \infty$ , para todo  $t$ , lo cual indica que la variabilidad del proceso es constante y finita.
- iii.  $\text{Cov}(Y_t, Y_{t-k}) = \gamma_k$ . Es decir la covarianza entre  $Y_t$  y  $Y_{t-k}$  solo depende de la distancia que hay entre  $t$  y  $t-k$ .

Si falta al menos una de las condiciones se dice que el proceso es no estacionario de segundo orden.

## 2.3 Proceso de Ruido Blanco

Sea  $\{Z_t\}$  una colección de variables aleatorias no correlacionadas con media cero y varianza constante  $\sigma^2$ , a la colección  $Z_0, Z_1, Z_2, \dots$ , se le conoce como proceso de ruido blanco.

Es importante comentar que el proceso de ruido blanco no requiere que las variables aleatorias sean independientes, ya que como es sabido, correlación cero no indica independencia de variables aleatorias, excepto cuando las variables aleatorias tienen distribución normal. Con este razonamiento, entonces, cualquier colección de variables aleatorias  $\text{iid} \sim (0, \sigma^2)$  es un proceso de ruido blanco pero lo contrario no es necesariamente cierto.

## 2.4 Función de Autocorrelación (ACF)

La autocorrelación de orden  $k$  para un proceso estocástico  $\{Y_t\}$ , denotada por  $\rho_k$ , es la correlación de las variables separadas  $k$  períodos, es decir:

$$\rho_k = \text{Corr}(Y_t, Y_{t-k}) = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{V(Y_t)}\sqrt{V(Y_{t-k})}}, \quad k \in \mathbb{Z}$$

y si el proceso es estacionario, entonces la autocorrelación de orden  $k$  es:

$$\rho_k = \frac{\gamma_k}{V(Y_t)} = \frac{\gamma_k}{\gamma_0}, \quad k \in \mathbb{Z}$$

A la función  $\rho_k$  para los diferentes valores de  $k$  se le llama función de autocorrelación poblacional y se denota como ACF. El gráfico de los valores de autocorrelación se llama correlograma.

Si en el correlograma los valores decaen rápidamente a cero, o más formalmente si  $\sum_k |\rho_k| < \infty$  entonces se dice que el proceso  $Y_t$  es estacionario.

## 2.5 Función de Autocorrelación Parcial (PACF)

La autocorrelación Parcial entre  $Y_t$  y  $Y_{t-k}$  se obtiene como la correlación que existe entre  $Y_t$  y  $Y_{t-k}$  después de eliminar el efecto de todas las variables aleatorias que están entre ellas, se denota por  $\phi_{kk}$  y puede calcularse a través de la FAC, en la forma:

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

.

.

.

# CAPÍTULO 3

## FORMULACIÓN GENERAL DE MODELOS LINEALES

La formulación general que presenta un modelo ARIMA de orden (p, d, q), es decir el modelo ARIMA (p, d, q) es la siguiente:

$$\phi_p(B) (1 - B)^d X_t = c + \theta_q(B) \varepsilon_t, \quad (1)$$

donde  $X_t$  es la variable de estudio,  $c$  una constante y  $\varepsilon_t$  es el término de error que sigue una distribución normal de media cero y varianza constante  $\gamma_\varepsilon^2$ . El término  $(1 - B)^d$  se aplica a la serie original para convertirla en estacionaria, y  $d$  corresponde al orden de integración  $\phi_p(B)$  y  $\theta_q(B)$  son polinomios de orden  $p$  y  $q$  que dependen del operador de retardo  $B$ .

El operador de retardos  $B$  está definido por la ecuación:

$$B^d X_t = X_{t-d}$$

El polinomio  $\phi_p(B)$  se define como:

$$\phi_p(B) = 1 - \sum_{j=1}^p \phi_j B^j \quad (2)$$

tal que  $\phi_p(B)X_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p}$  y donde  $\phi_i$  ( $i = 1, \dots, p$ ) son los coeficientes del polinomio  $\phi_p(B)$ ,  $p$  es el número de términos del polinomio  $\phi_p(B)$  y es el orden correspondiente a la parte AR del modelo ARIMA.

El polinomio  $\theta_q(B)$  se define como:

$$\theta_q(B) = 1 - \sum_{j=1}^q \theta_j B^j \quad (3)$$

Donde  $\theta_q(B)\varepsilon_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$  y donde  $\theta_j$  ( $j = 1, \dots, q$ ) son los coeficientes del polinomio  $\theta_q(B)$ ,  $q$  es el número de términos del polinomio  $\theta_q(B)$  y el orden correspondiente a la parte MA del modelo ARIMA.

Por lo tanto si sustituimos (2) y (3) en la expresión (1) se obtiene:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d X_t = c + (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t$$

En conclusión, el modelo ARIMA está compuesto de tres partes: una parte AR de orden  $p$ , una parte Integrada de orden  $d$  y una parte MA de orden  $q$ .

El número de términos para los polinomios  $\phi_p(B)$  y  $\theta_q(B)$ , es decir, los órdenes de la parte AR y MA respectivamente, así como el orden de la parte I del modelo Arima se identifican usando la metodología Box Jenkins

### 3.2 Pasos en la construcción de modelos ARIMA

#### Paso 1: Identificación de los términos del modelo

Primero identificamos las principales características de la serie temporal: como su comportamiento estacionario y la presencia de estacionalidad de los datos.

Para la corrección de la no estacionariedad se pueden realizar dos tipos de transformaciones:

- Para estabilizar la varianza normalmente se usan transformaciones de Box Cox; como el logaritmo natural, raíz cuadrada, etc.
- Para estabilizar la media se toman diferenciaciones de orden  $d$  del tipo :

$$(1 - B)^d X_t$$

Existe estacionalidad en los datos cuando los datos que componen la serie presentan un comportamiento cíclico o periódico. Por ejemplo la serie de precios presenta estacionalidad semanal, un día puede ser parecido al mismo día pero de la semana anterior; es decir los lunes tienden a ser parecidos a los lunes, los martes similares a los martes, y así sucesivamente.

Si los datos presentan estacionalidad la formulación del modelo Arima incluye el modelo:

$$\phi_p(B) \Phi_P(B^s) Y_t = \theta_q(B) \Theta_Q(B^s) Z_t$$

donde  $s$  representa el orden de estacionalidad que presentan los datos,  $d$  corresponde a la parte integrada del modelo Arima estacional. Normalmente  $d$  toma los valores 1 o 2, y  $\phi(B^s)$  y  $\theta(B^s)$  son polinomios, AR y MA de orden  $P$  y  $Q$  y que sólo dependen de los retardos  $B^s$ .

#### Paso 2: Estimación de los parámetros del modelo

Una vez identificados los términos que contiene el modelo se estiman los parámetros que lo constituyen.

La estimación de los parámetros del modelo se puede hacer a través de diferentes métodos. El método más utilizado es el de la máxima verosimilitud, aunque en los métodos autorregresivos una estimación de fácil utilización es el método de los momentos.

La maximización de la función de verosimilitud es no lineal en el sentido de que la función a maximizar no es una función cuadrática de los parámetros desconocidos. Esta maximización es por lo tanto realizada numéricamente. Por ello la convergencia al máximo será más rápida si se parte de un valor inicial de los parámetros próximo al de convergencia. Hay distintos métodos para el cálculo de estos valores iniciales, dos de ellos para el caso autorregresivo (método de Yule-Walker y algoritmo de Burg) y otros dos para un caso general (algoritmo de las innovaciones y algoritmo de Hannan-Rissanen).

### Paso 3: Validación del modelo

1. Significancia estadística de los parámetros
2. Verificación de estacionariedad e invertibilidad
3. Análisis de residuos:

Para asegurar la validez e idoneidad del modelo y la efectividad de las predicciones, los residuos estimados se deben comportar como un ruido blanco.

Para comprobar que los residuos estimados son de ruido blanco se debe tener en cuenta:

- Representamos FAC y FACP para los residuos: si los residuos son de ruido blanco, tanto en la FAC como en las FACP de estos residuos no debe aparecer ningún valor significativo; es decir los valores de estas funciones deben ser muy pequeños y estar dentro de las bandas de confianza.

$$\left[ -\frac{1,96}{\sqrt{T}}, \frac{1,96}{\sqrt{T}} \right]$$

Estas son bandas de asintóticas al 95% de confianza, donde T es el número de valores de la serie  $X_t$ .



- Test de Ljung-Box: este test indica si existe dependencia entre los m primeros residuos estimados, es decir, si estos residuos presentan correlación no nula.

El estadístico de Ljung-Box se define como:

$$Q = T(T+2) \sum_{h=1}^m \frac{\rho_h^2}{T-h}$$

Donde  $\hat{\rho}_h$  es el coeficiente de autocorrelación de los residuos estimados. T es el número de valores de la serie  $X_t$  y r es el número de parámetros estimados.

Este estadístico, Q, se distribuye como una Chi-cuadrado con un número de grados de libertad igual a:  $m - r - 1$ .

En la mayoría de los casos es suficiente con presentar la FAC y FACP, ya que si no presentan valores significativos, el valor estadístico de Q será pequeño, y por tanto se puede considerar que existe independencia entre los residuos.

#### Paso 4: Predicción

Después de obtener el modelo y comprobar su validez, se puede proceder a predecir.

La predicción óptima de  $X_{T+k}$ ,  $\hat{X}_{T+k}$ , es el valor esperado de  $X_{T+k}$  condicionado a que se conoce  $X_1, X_2, \dots, X_T$ , es decir, la esperanza condicionada de  $X_{T+k}$  conocido  $X_1, X_2, \dots, X_T$ . De forma análoga se procede con los residuos.

Por lo tanto:

$$\hat{x}_t(k) = E[X_{T+k} | X_T \dots X_1]$$

$$\hat{\varepsilon}_t(k) = E[\varepsilon_{T+k} | X_T \dots X_1]$$

Donde T representa el origen de la predicción y k el horizonte de la misma. Las fórmulas correspondientes a las predicciones que se quieren obtener, según el modelo estimado, son:

$$X_{T+k} = \hat{\phi}_1(X_{T+k-1}) + \dots + \hat{\phi}_p(X_{T+k-p}) + \hat{c} + \varepsilon_{T+k} + \hat{\theta}_1\varepsilon_{T+k-1} + \dots + \hat{\theta}_q\varepsilon_{T+k-q}$$

Tomando esperanzas condicionadas en la expresión anterior, la ecuación de predicción para el modelo Arima estimado es la siguiente:

$$X_{T+k} = \hat{\phi}_1(X_{T+k-1}) + \dots + \hat{\phi}_p(X_{T+k-p}) + \hat{c} + \varepsilon_T(k) + \hat{\theta}_1\hat{\varepsilon}_T(k-1) + \dots + \hat{\theta}_q\hat{\varepsilon}_T(k-q)$$

Donde:

$\hat{X}_T(j) = X_{T+j}$  si  $j \leq 0$ , es el valor de la serie  $X_t$  en el tiempo  $T+j$

$\hat{X}_T(j) = \hat{X}_{T+j}$  si  $j > 0$ , es la predicción obtenida para la serie  $X_t$  en el tiempo  $T+j$

$\hat{\varepsilon}_T(j) = \hat{\varepsilon}_{T+j}$ , si  $j \leq 0$  es el valor de la serie  $\hat{\varepsilon}_T$  en el tiempo  $T+j$

$\hat{\varepsilon}_T(j) = 0$ , si  $j > 0$

### 3.3 Formulación general de modelos no lineales

#### 3.3.1 Modelos Arch

Los modelos Arch o modelos autorregresivos con heterocedasticidad condicionada, fueron introducidos por Engle (1982), con el objetivo de estimar la variación de la inflación. La idea básica es que el retorno  $\gamma_t = \ln(x_t) - \ln(x_{t-1})$  no está correlacionado serialmente pero presenta una volatilidad (varianza condicional) que depende de los retornos pasados por medio de una función cuadrática.

Un modelo Arch (r) está definido por:

$$X_t = \sqrt{h_t} \varepsilon_t,$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2,$$

Donde  $\varepsilon_t$  es una secuencia de variables aleatorias independientes e idénticamente distribuidas (i.i.d) con media cero varianza uno,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i > 0$ .

En la práctica, usualmente suponemos  $\varepsilon_t \sim N(0,1)$  como  $\varepsilon_t \sim t_\nu$  (distribución t Student con  $\nu$  grados de libertad).

Los coeficientes  $\alpha_i$  deben satisfacer ciertas condiciones, dependiendo del tipo de imposición que colocamos sobre el proceso.

### 3.3.1.2 Modelo Arch (1)

El modelo Arch (1) tiene la forma:

$$X_t = \sqrt{h_t} \varepsilon_t,$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2,$$

con  $\alpha_0 > 0$ ,  $\alpha_1 \geq 0$ .

Calculemos la media, la varianza y autocovarianza incondicionales de la serie.

$$i. E(X_t) = E\{E(X_t | \mathcal{F}_{t-1})\} = 0$$

$$ii. Var(X_t) = E(X_t^2) = E\{E(X_t^2 | \mathcal{F}_{t-1})\} = E(\alpha_0 + \alpha_1 X_{t-1}^2) = \alpha_0 + \alpha_1 E(X_{t-1}^2).$$

Si el proceso  $\{X_t\}$  es estacionario de segundo orden, entonces para todo  $t$ ,  $E(X_t^2) = E(X_{t-1}^2) = Var(X_t)$ , entonces:

$$Var(X_t) = \frac{\alpha_0}{1-\alpha_1}$$

Y como  $Var(X_t) > 0$ , entonces  $0 \leq \alpha_1 < 1$ .

$$iii. Cov(X_{t+k}, X_t) = E(X_{t+k} \cdot X_t), \quad k \geq 1$$

$$= E[E(X_{t+k}, X_t) | \mathcal{F}_{t+k-1}] = E[X_t E(X_{t+k} | \mathcal{F}_{t+k-1})]$$

$$= E[X_t E(\sqrt{h_{t+k}} \varepsilon_{t+k} | \mathcal{F}_{t+k-1})] = 0, \text{ pues } E(\varepsilon_{t+k} | \mathcal{F}_{t+k-1}) = 0$$

$$\sigma_X(k) = 0, \quad \forall k \geq 1$$

Indicando que  $X_t$  es una secuencia de variables aleatorias no correlacionadas (ruido blanco).

Para calcular la curtosis, es necesario calcular el momento de cuarto orden de  $X_t$ . Suponemos que  $\varepsilon_t$  son normales. Tenemos:

$$E(X_t^4 | \mathcal{F}_{t-1}) = E(h_t^2 \varepsilon_t^4 | \mathcal{F}_{t-1}) = 3(\alpha_0 + \alpha_1 X_{t-1}^2)^2$$

Pues  $E(\varepsilon_t^4) = 3$  de donde:

$$E(X^4) = 3E(\alpha_0 + \alpha_1 X_{t-1}^2)^2 = 3(\alpha_0^2 + 2\alpha_0\alpha_1 X_{t-1}^2 + \alpha_1^2 X_{t-1}^4)$$

Admitiendo que el proceso es estacionario de cuarto orden, el momento de cuarto orden puede ser escrito  $\mu_4 = E(X_t^4)$  y tenemos que:

$$\mu_4 = 3(\alpha_0^2 + 2\alpha_0\alpha_1 \text{Var}(X_t) + \alpha_1^2 \mu_4)$$

$$= 3[\alpha_0^2 + 2\alpha_0\alpha_1(\frac{\alpha_0}{1-\alpha_1}) + \alpha_1^2 \mu_4]$$

$$= 3\alpha_0^2(1 + 2(\frac{\alpha_1}{1-\alpha_1})) + 3\alpha_1^2 \mu_4$$

de aquí obtenemos finalmente que:

$$\mu_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$$

Suponiendo que los momentos de cuarto orden sean finitos y positivos debemos tener en cuenta  $1 - 3\alpha_1^2 > 0$ , o sea que:  $0 \leq \alpha_1^2 < \frac{1}{3}$ . Por lo tanto, cuanto más restricciones colocamos a el proceso de retornos, más restricciones tenemos para los coeficientes del modelo.

La curtosis  $X_t$ , estará dada por:

$$K = \frac{\mu_4}{[\text{Var}X_t]^2} = 3 \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \frac{(1-\alpha_1)^2}{\alpha_0^2} = 3 \frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3$$

Vemos pues, que si admitimos que  $X_t$  sigue un modelo ARCH, las colas serán más pesadas que las de la Normal, que es una propiedad ventajosa del modelo. Por otro lado, una desventaja del modelo es que trata los retornos positivos y negativos de forma similar, ya que los cuadrados de los retornos entran en una fórmula de volatilidad. En la práctica se sabe que la volatilidad responde de modo diferente en retornos positivos y negativos.

En resumen:

- i. Las esperanzas marginal y condicional son iguales a cero.
- ii. La varianza marginal es constante
- iii. La varianza condicional depende de los valores que haya tomado  $e_{t-1}^2$  luego no es constante.
- iv. La distribución marginal del proceso ARCH (1) tiene una forma desconocida.

### Modelo Arch ( r )

El modelo anterior puede generalizarse permitiendo una dependencia de la varianza condicional con r retardos. De manera que el modelo Arch ( r ) para  $X_t = \sigma_t \varepsilon_t$ , la varianza condicional:

$$h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2$$

Donde  $\alpha_0 > 0$ , (corresponde a la mínima varianza condicional observada) y  $0 \leq \alpha_1 \leq 1$  (es una condición necesaria y suficiente para la existencia de varianza incondicional y la condicional).

En este proceso las posibilidades de rachas de alta volatilidad dependen de los r últimos valores. La varianza marginal:

$$\text{Var} (X_t) = E (X_t^2) = E [E (X_t^2 | X_{t-1})] = \alpha_0 + \sum \alpha_i E (X_{t-i}^2)$$

Y Por lo tanto:

$$\text{Var} (X_t) = \frac{\alpha_0}{1 - \alpha_1 - \dots - \alpha_r}$$

De donde:

$$\sum_{i=1}^r \alpha_i < 1$$

Si introducimos  $\vartheta_t = X_t^2 - h_t$ , como en el caso del proceso Arch (1), será un proceso de ruido blanco, formado por variables estacionarias no correlacionadas de media cero y varianza marginal constante, podemos expresar la dependencia de los cuadrados de las observaciones como un proceso AR( r ):

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2 + \vartheta_t$$

Estas variables no son independientes entre sí ni de los regresores, ya que la positividad de  $X_t^2$  exige que:

$$\vartheta_t > - (\alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_r X_{t-r}^2)$$

Así en un modelo Arch ( r ) se verifica que :

i. Es un proceso de ruido blanco pero no es independiente y no está idénticamente distribuido.

- ii. Las esperanzas condicional y no condicional son iguales a cero.
- iii. La varianza no condicional es constante
- iv. La varianza condicional depende de  $X_{t-1}$ ,  $X_{t-2}$ , ...,  $X_{t-r}$ , luego no es constante.

### 3.3.2 Modelo Garch

Es un rasgo común a muchas de las primeras aplicaciones empíricas de los modelos Arch es que requieren un gran número de parámetros autorregresivos y, para representar adecuadamente el comportamiento dinámico de la varianza, se imponía una estructura fija de retardos. Con el fin de flexibilizar estas restricciones Bollerslev (1986) propuso un modelo Arch generalizado o Garch.

La generalización del modelo Arch al modelo Garch tiene gran similitud con la extensión de los procesos autorregresivos a los AR de medias móviles, ARMA, permitiendo una representación más parsimoniosa de la volatilidad. Bollerslev considera que la varianza,  $h_t^2$ , no sólo depende de las observaciones pasadas de  $X_t$ , depende también de su propio pasado. Esta dependencia se expresa incluyendo cierto número de retardos  $p$  de  $h_t^2$ , de modo que la varianza condicional se define como:

$$h_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}^2$$

donde  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, p$  aunque estas restricciones se establecen para garantizar que la varianza sea positiva, Nelson y Cao (1992) demuestran posteriormente que la positividad de la varianza está asegurada bajo condiciones más débiles. En concreto demuestran que si el modelo Garch admite una representación Arch ( $\infty$ ), es suficiente exigir que los coeficientes del polinomio de retardos en dicha representación sean todos positivos. El nuevo modelo se llama Garch ( $p, r$ ) y se reduce al Arch ( $r$ ) cuando  $p = 0$ . Bollerslev establece las condiciones de estacionariedad, probando que  $X_t$  es débilmente estacionario con  $E(X_t) = 0$ ,  $Var(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^r \alpha_i - \sum_{j=1}^p \beta_j}$ .

#### Modelo Garch (1,1)

Muchos trabajos con series financieras, muestran que el más sencillo de los modelos Garch, el Garch (1,1), es suficiente para modelizar con éxito los cambios temporales en la varianza condicional, incluso sobre los períodos muestrales

largos. El modelo Garch (1,1) se obtiene cuando  $p = r = 1$ , de modo que la varianza condicional queda:

$$h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1}^2$$

Con  $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$ . Si  $\alpha_1 + \beta_1 < 1$ , la serie  $X_t$  tiene varianza finita, y por ser una diferencia martingala<sup>1</sup>, es ruido blanco, de media cero y varianza:

$$\text{Var}(X_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

Además, Bollerslev prueba que si:  $(\beta_1 + \alpha_1)^2 + 2\alpha_1^2 < 1$ , el momento de cuarto orden de  $X_t$  existe y es finito, y la curtosis de  $X_t$  es:

$$K = \frac{E(X_t^4)}{[E(X_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2}$$

Cuando  $\alpha_1 + \beta_1 < 1$ , este valor es mayor que 3 y, por tanto el proceso Garch (1,1) estacionario es leptocurtica, propiedad que comparte con el Arch (1).

### 3.3.3 Modelo Igarch

En las aplicaciones de modelos Garch (1,1) a series financieras, es casi sistemática la obtención de un valor estimado de  $\alpha_1 + \beta_1$  prácticamente igual a uno, en especial si la frecuencia de observación es alta. Teniendo en cuenta la forma de la función de autocorrelación de  $X_t^2$ , un valor de  $\alpha_1 + \beta_1$  próximo a uno significa que dicha función apenas decrece, indicando que los cambios de la varianza condicional son relativamente lentos y, por tanto, los shocks (cambios bruscos) en la volatilidad persisten. Esta propiedad es interesante porque refleja precisamente una de las características de las series financieras: aunque la serie original está es no autocorrelacionada, existe correlación en la serie de los cuadrados y, además, estas correlaciones decrecen lentamente, mostrando valores significativamente distintos de cero incluso para retardos altos.

Los resultados anteriores justifica el interés de un modelo Garch (1,1) en el que se imponga la condición  $\alpha_1 + \beta_1 = 1$ . El modelo resultante denominado Garch

---

<sup>1</sup>Sea  $X_0, X_1, X_2, \dots$  una cadena de Markov. La cadena es una Martingala si para todo  $n = 0, 1, 2, \dots$ , tenemos que:  $E(X_{n+1} - X_n | X_n) = 0$ . Es decir en promedio el valor de la cadena no varía, sin importar cuál sea el valor  $X_n$  en un momento dado

Integrado, fue propuesto por Engle y Bollerslev (1986) y en él la ecuación para la varianza condicionada es:

$$h_t^2 = \alpha_0 + h_{t-1}^2 + \alpha_1(X_{t-1}^2 - h_{t-1}^2)$$

El modelo no es débilmente estacionario, porque su varianza marginal no está definida.

El Modelo es:

$$X_t = \sqrt{h_t} \varepsilon_t$$

$$h_t = \alpha_0 + \beta_1 h_{t-1} + (1 - \beta_1) X_{t-1}^2$$

### 3.3.4 Modelo Egarch

Nelson observó ciertas limitaciones en los modelos Garch:

- Las condiciones impuestas sobre los parámetros para asegurar que  $h_t^2$  no sea negativo no se satisfacen en algunas aplicaciones empíricas.
- El modelo Garch es incapaz de modelizar una respuesta asimétrica de la volatilidad ante las subidas y bajadas de la serie.

Con el fin de solventar estas deficiencias, Nelson propuso un nuevo modelo Garch exponencial o Egarch.

El modelo Egarch garantiza la no negatividad de la varianza condicional, formulando la ecuación de la volatilidad en términos del logaritmo de  $h_t^2$ , mediante una representación lineal del tipo:

Modelo Egarch (1, 1)

$$X_t = \sqrt{h_t} \varepsilon_t$$

$$\ln(h_t) = \alpha_0 + \alpha_1 g(\varepsilon_{t-1}) + \beta_1 \ln(h_{t-1})$$

donde  $g(\varepsilon_t) = \lambda_1 \varepsilon_t + \lambda_2 (|\varepsilon_t| - E|\varepsilon_t|)$ : curva de impacto de información. A través de esta función  $g$ , que depende del signo y de la magnitud de  $\varepsilon_t$ , el modelo Egarch puede capturar una respuesta asimétrica de la volatilidad ante innovaciones de distinto signo, permitiendo así modelizar un efecto contrastado empíricamente en series financieras: las malas noticias (rendimientos negativos) provocan mayor aumento de la volatilidad que las buenas noticias (rendimientos positivos).



### 3.3.5 Modelo Tarch (1,1)

O modelo TARCH ( “ Threshold ARCH “) es un caso particular del modelo no lineal ARCH y la volatilidad que sigue la forma funcional:

$$h_t^\gamma = \alpha_0 + \alpha_1 g^\gamma(\varepsilon_{t-1}) + \beta_1 h_{t-1}^\gamma, \text{ en la cual:}$$

$$g^\gamma(\varepsilon_t) = \theta I_{(\varepsilon_t > 0)} |\varepsilon_t|^\gamma + (1 - \theta) I_{\{\varepsilon_t \leq 0\}} |\varepsilon_t|^\gamma$$

Para  $\gamma = 1$  tenemos el modelo de Zakoian (1994) e para  $\gamma = 2$  el modelo GJR (de Glosten, Jagannathan and Runkle, 1993).

En Eviews usamos la fórmula:

$$h_t = w + \alpha X_{t-1}^2 + \gamma X_{t-1}^2 d_{t-1} + \beta h_{t-1},$$

$$d = \begin{cases} 1 & \text{si } X_t < 0 \\ 0 & \text{si } X_t \geq 0 \end{cases}$$

d : variable Dummy, solo puede asumir los valores de 0 y 1.

### 3.4 Construcción de Modelos no lineales

#### Paso 1: Identificación de los términos del modelo

La identificación de los modelos Arch y Garch, se efectúa después de ajustar un modelo Arima a la serie. Si existen efectos Arch, los residuos del modelo Arima estarán incorrelados pero no serán independientes y este efecto será visible en la función de autocorrelación de los residuos al cuadrado, que mostrará correlación serial. Además, si calculamos los coeficientes de autocorrelación parcial de los residuos al cuadrado y el modelo para los residuos es Arch puro, el número de términos distintos de cero nos indicará, aproximadamente, el orden del proceso.

## Paso 2: Estimación de parámetros del Modelo

Estimación de parámetros del Modelo Arch: Se construye la función de verosimilitud utilizando la descomposición del error de predicción. Maximizando esta función se obtienen los estimadores máximo verosímiles. Como es habitual en modelos de series temporales, la función de verosimilitud se construye como el producto de las densidades condicionadas. Asumiendo que las perturbaciones  $\varepsilon_t$  en  $X_t = h_t \varepsilon_t$ , son variables aleatorias independientes e idénticamente distribuidas con distribución  $N(0,1)$ , los modelos Arch son condicionalmente Gaussianos y la distribución condicionada  $X_t/X_{t-1}$  es  $N(0, h_t^2)$ . Por tanto la expresión del logaritmo de la función de verosimilitud resulta ser:

$$\ln L(\theta) = \sum \ln f(X_t/\theta, X_{t-1})$$

Donde  $\theta$  es el vector de parámetros desconocidos del modelo  $f(X_t/\theta, X_{t-1})$  denota la densidad condicionada de  $X_t$  dadas las observaciones previas hasta el instante  $t-1$ .

Para facilitar cálculos vamos a considerar un modelo Arch (1), su función de verosimilitud es:

$$f(x_1, \dots, x_t/\alpha_0, \alpha_1) = f(X_1)f(X_2/X_1)\dots f(X_t/X_{t-1})$$

Donde las funciones de densidad  $f(X_t/X_{t-1})$  son normales. Como  $X_t = h_t \varepsilon_t$ , si consideramos  $X_{t-1}$ , el valor de  $h_t$  es una constante y la única variable es  $\varepsilon_t$  que tiene distribución normal. La media condicionada de la distribución  $f(X_t/X_{t-1})$  es cero, y la varianza  $h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$ . El logaritmo de la función de verosimilitud condicionada de  $X_2, \dots, X_T$  dado  $\alpha_0, \alpha_1$  y  $X_1$  es:

$$L(x_2, \dots, x_t/\alpha_0, \alpha_1) = -\frac{t-1}{2} \ln 2\pi - \frac{1}{2} \sum_{t=2}^T \ln(\alpha_0 + \alpha_1 x_{t-1}^2) - \frac{1}{2} \sum_{t=2}^T \frac{x_t^2}{\alpha_0 + \alpha_1 x_{t-1}^2}$$

Derivando con respecto a los parámetros, llamando  $\widehat{h_t^2} = \widehat{\alpha_0} + \widehat{\alpha_1} X_{t-1}^2$  e igualando a cero se obtienen las ecuaciones:

$$\sum \frac{x_t^2}{\widehat{h_t^4}} = \sum \frac{1}{\widehat{h_t^2}}$$

$$\sum \frac{X_{t-1}^2}{h_t^2} = \sum \frac{X_{t-1}^2 X_t^2}{h_t^4}$$

Multiplicamos y dividimos el primer miembro por  $\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2$ :

$$\widehat{\alpha}_0 \sum \frac{1}{\sigma_t^2} + \widehat{\alpha}_1 \sum \frac{X_{t-1}^2}{\sigma_t^4} = \sum \frac{x_t^2}{\sigma_t^4}$$

$$\widehat{\alpha}_0 \sum \frac{x_{t-1}^2}{\sigma_t^4} + \widehat{\alpha}_1 \sum \frac{x_{t-1}^4}{\sigma_t^4} = \sum \frac{x_{t-1}^2 x_t^2}{\sigma_t^4}$$

Son las ecuaciones mínimos cuadrados para obtener los parámetros del modelo. Resolviendo este sistema obtenemos la estimación de  $\widehat{\alpha}_0$  y  $\widehat{\alpha}_1$ .

### Paso 3: Diagnóstico

Vamos a llamar a  $e_t$  a los residuos del modelo ARIMA y  $\widehat{\sigma}_t^2$  a las varianzas condicionadas, los residuos estandarizados  $\frac{e_t}{\widehat{\sigma}_t}$ , deben seguir un proceso de ruido blanco normal y podemos aplicarles los contrastes para los procesos Arima. Sus cuadrados no deben mostrar dependencia, este se puede comprobar con los contrastes de autocorrelación de los cuadrados: Durbin Watson, Wallis, h-Durbin, Breusch-Godfrey y Cochrane-Orcutt.

Box y Jenkins sugirieron un número considerable de tests para verificar si el modelo elegido se ajusta correctamente al conjunto de datos dados. Uno de ellos, conocido como sobreparametrización, consiste en ajustar un modelo de orden superior al elegido y comprobar si los parámetros son significativamente distintos de cero.

- De otro lado si el modelo se aproxima satisfactoriamente a la serie observada, los residuos deben tender a comportarse como ruido blanco, lo que se comprobaría mediante funciones de autocorrelación de residuos (ACF, ACFP). Dichas funciones de autocorrelación deben ser nulas en todo su recorrido, excepto en el cero.
- Si el modelo no se aproxima satisfactoriamente a la serie observada los residuos se comportarían como un ruido autocorrelacionado. Por ello deben, deben emplearse contrastes como el de Durbin-Watson (para la autocorrelación de primer orden).

- Box y Pierce proponen el estadístico  $Q = \sum_{k=1}^m r_k^2$ , donde  $r_k = \frac{\sum_{t=k+1}^n \varepsilon_t \varepsilon_{t-k}}{\sum_{t=1}^n \varepsilon_t^2}$ , siendo  $\varepsilon_t \equiv$  residuos estimados y  $n$  el número de observaciones. Bajo el supuesto de que  $m$  es suficientemente grande, Box y Pierce demuestran que el estadístico  $Q$  se distribuye como una Chi-cuadrado. Rechazándose la hipótesis de que los residuos son ruido blanco para valores de  $Q$  muy altos. Más concretamente, se halla la región crítica para un nivel de significación  $\alpha$ , calculando un valor  $I$  que cumpla  $P(Q > I) = \alpha$ . Cuando el valor de  $Q$  cae dentro de la región crítica se rechaza la hipótesis nula de que los residuos son un ruido blanco. Si cae fuera de la región crítica se acepta la hipótesis nula.
- Un diagnóstico completo surge de la inspección gráfica de los residuos. Si los residuos provienen de un proceso de ruido blanco, deben ser incorrelacionados entre sí, lo que les hará alternar en signo, sin ningún criterio obvio. Por el contrario la racha de residuos consecutivos de un mismo signo son, en general, un indicativo de mala especificación del modelo, bien por ser una indicación de autocorrelación de los residuos o por indicar no estacionariedad de los mismos. Si el gráfico  $(t, \varepsilon_t)$  tiene una tendencia conocida, puede haber heterocedasticidad de los residuos.

En la evaluación de los modelos, hay que tener en cuenta la posible confusión entre valores atípicos y heterocedasticidad condicional. Los valores atípicos pueden interpretarse como un aumento en la varianza en ese instante y especialmente si aparecen rachas, puede confundirse con efectos de heterocedasticidad condicional.

Por otro lado, la serie que sigue un modelo Arch puede mostrar muchos valores atípicos si se analiza como si tuviese varianza constante y sigue un proceso ARMA. Por lo tanto es importante diferenciar ambos fenómenos.

En la práctica, se suele limpiar la serie inicialmente de las observaciones que presentan residuos tan grandes que no pueden ser debidas a heterocedasticidad condicional y que son muy probablemente valores atípicos. Una regla simple y efectiva es que es muy poco probable que la heterocedasticidad condicional pueda generar observaciones con residuos mayores de siete desviaciones típicas, y considerar estos datos atípicos.

# CAPÍTULO 4

## ESTUDIO DE LA SERIE DE TIEMPO: PRECIO EXTERNO DEL CAFÉ COLOMBIANO

Para el análisis de la serie de tiempo: precio externo del café colombiano, consideramos una serie que va desde el primero de enero de 1980 hasta diciembre del 2015.

### 4.1 Análisis exploratorio de la serie $Z_t$ : precio externo del café colombiano

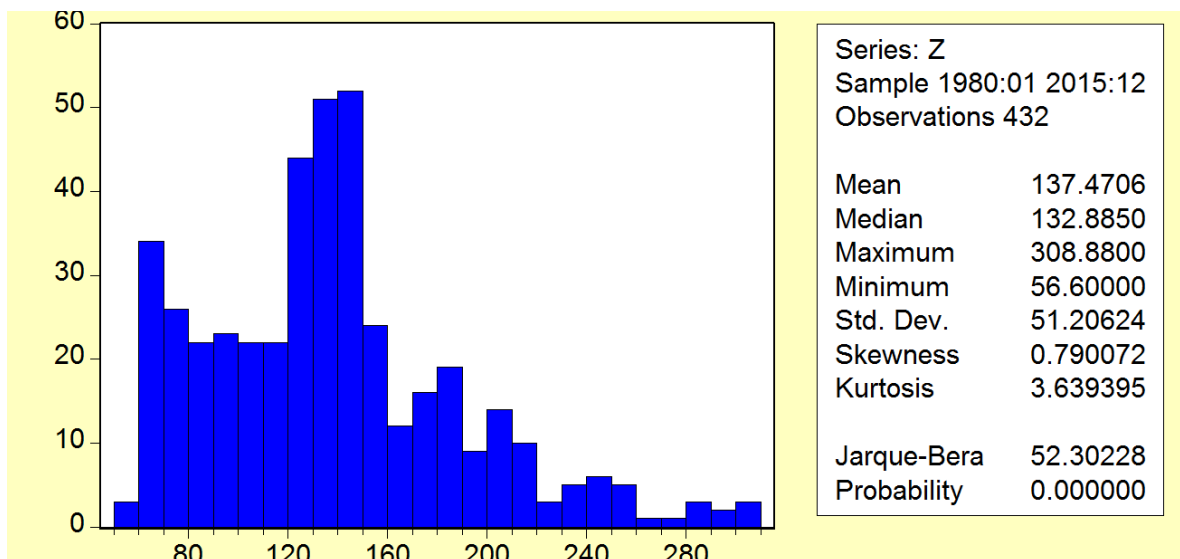


Figura 4.1: Histograma de la serie precio externo del café colombiano

Observamos que la distribución del precio externo del café colombiano es leptocúrtica (una de las principales características de las series financieras) ya que el coeficiente de kurtosis es mayor que el coeficiente que se obtiene de una distribución Normal. Esta es una prueba clara de que existe un mayor número de valores extremos de lo que se esperaría de una distribución normal estándar. El valor  $-p$  asociado al estadístico de prueba de Jarque Bera muestra que el precio externo del café colombiano no sigue una distribución normal.

Otra forma gráfica de verificar si la distribución de la serie de tiempo precio externo del café colombiano sigue una distribución normal es por medio del gráfico Quantile-Quantile. La figura presenta una forma de S alargada, lo que indica que la serie no sigue una distribución Normal y, además presenta valores extremos.

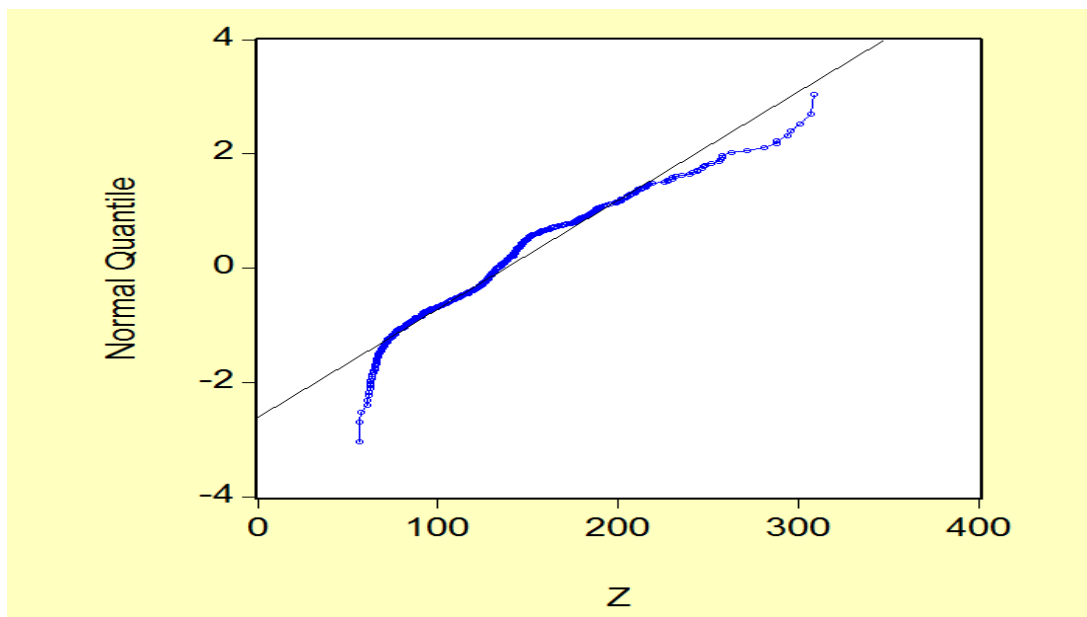


Figura 4.1.1: Gráfico Quantile-Quantile de la serie precio externo del café colombiano

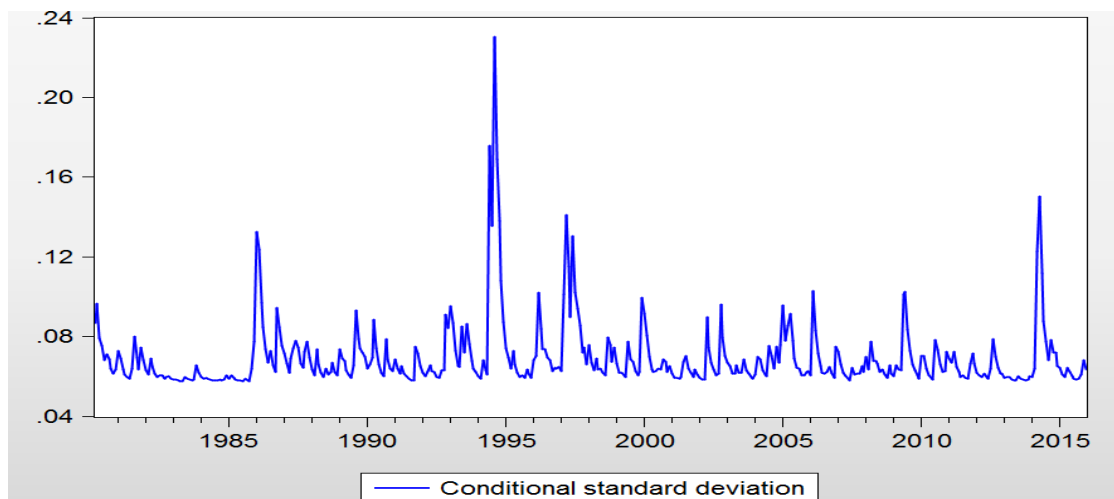


Figura 4.1.2: gráfico de desviación condicional estándar de la serie precio externo del café colombiano

La figura muestra que la serie precio externo del café colombiano presenta agrupamiento de volatilidad.

La figura 4.1.2 muestra el correlograma de la serie de tiempo: precio externo del café colombiano.

Date: 02/08/16 Time: 15:56  
Sample: 1980:01 2015:12  
Included observations: 432

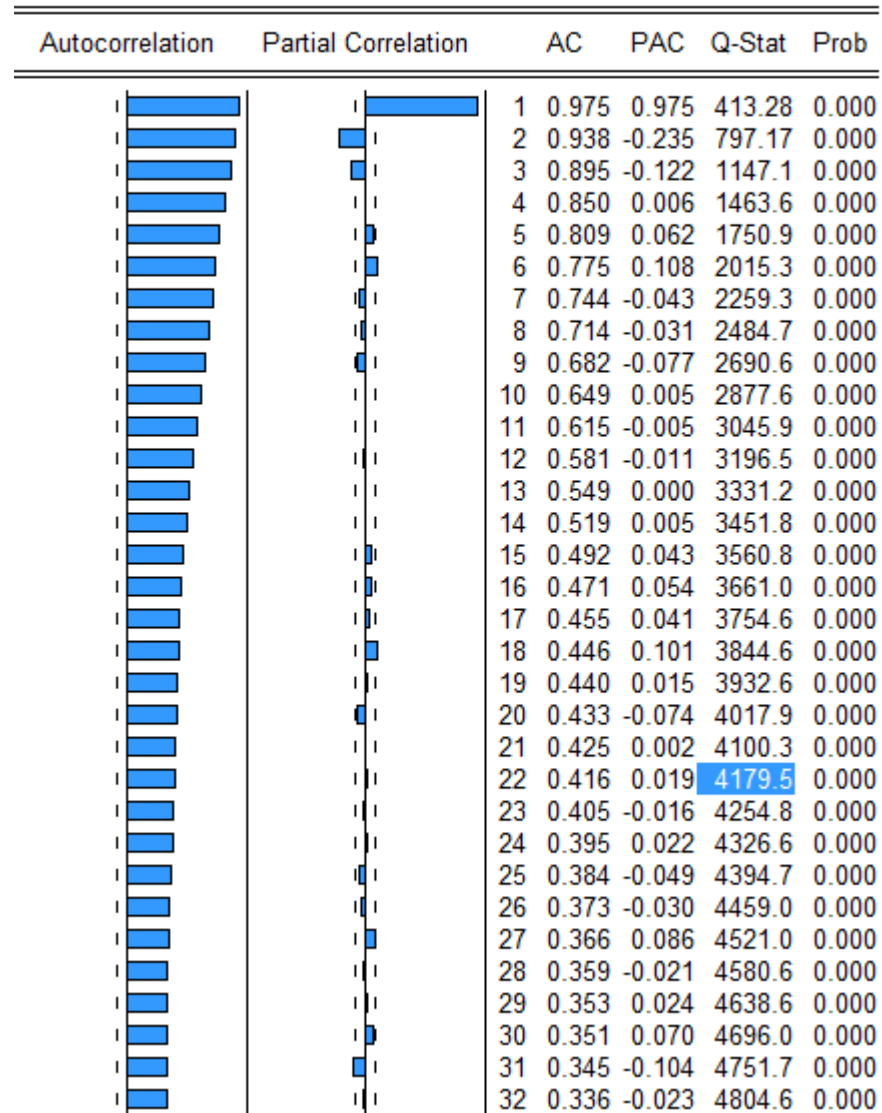


Figura 4.1.3: Correlograma de la serie precio externo del café colombiano

La serie no es estacionaria, además podríamos especular que tiene algo que ver con los modelos AR ( 1 ) .

Según El test de Dickey Fuller aplicado a la serie  $Z_t$  (precio externo del café colombiano) se concluye que la serie no es estacionaria.

Dependent Variable: Z  
 Method: Least Squares  
 Date: 02/08/16 Time: 16:15  
 Sample(adjusted): 1980:02 2015:12  
 Included observations: 431 after adjusting endpoints  
 Convergence achieved after 2 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.996458	0.003746	266.0352	0.0000
R-squared	0.950443	Mean dependent var		137.3935
Adjusted R-squared	0.950443	S.D. dependent var		51.24069
S.E. of regression	11.40691	Akaike info criterion		7.708634
Sum squared resid	55950.59	Schwarz criterion		7.718068
Log likelihood	-1660.211	Durbin-Watson stat		1.546006
Inverted AR Roots	1.00			

Tabla 4.1: Modelo AR (1) para la serie  $Z_t$

Aunque la serie no es estacionaria, tenemos la sospecha de que tiene alguna dependencia con los modelos autorregresivos, además el coeficiente AR de 0,99 que tiende a uno, nos dice que la serie es explosiva y probablemente tenga presente efectos Arch.

ADF Test Statistic	-3.045635	1% Critical Value*	-3.4478
		5% Critical Value	-2.8686
		10% Critical Value	-2.5705

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(Z)  
 Method: Least Squares  
 Date: 02/08/16 Time: 16:01  
 Sample(adjusted): 1980:06 2015:12  
 Included observations: 427 after adjusting endpoints

Tabla 4.2: Test de Dickey Fuller de la serie  $Z_t$

al 1% de significancia se acepta la hipótesis nula  $\phi^* = 0 \rightarrow \phi = 1$   $\phi^* = \phi^{-1}$

El test de Dickey Fuller nos muestra que la serie precio externo del café colombiano no es estacionaria.



Comparación de la serie precio externo del café colombiano ( $Z_t$ ) con la serie de los retornos, que es la serie pero después de haber aplicado logaritmo y una diferencia  $\gamma_t$  de orden 1.

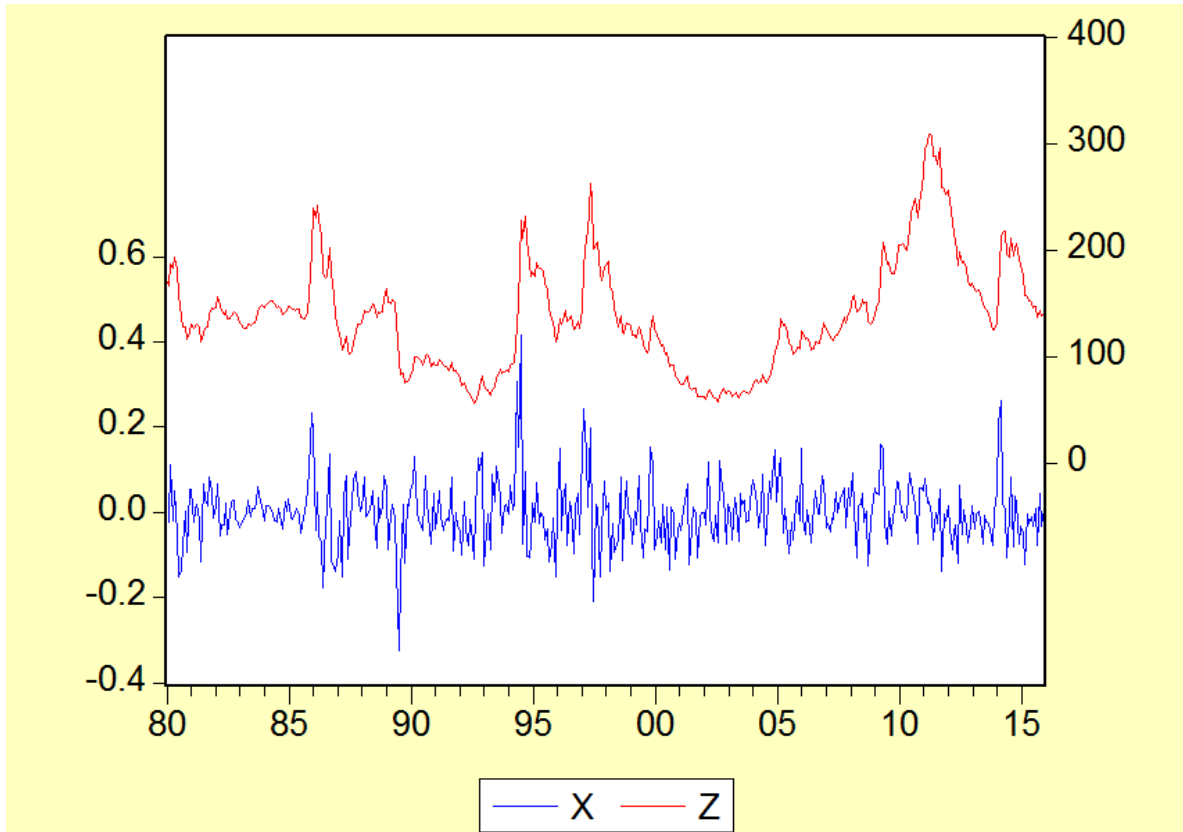


Figura 4.1.4: serie del precio externo del café colombiano  $Z_t$  y la serie de los retornos  $\gamma_t$ .

#### 4.2 Estudio de la volatilidad de la serie precio externo del café colombiano

Observamos que la serie presenta mucha volatilidad, por ejemplo en los años 1986, 1990, 1995, 1997, 1998, 2014. Lo cual nos hace pensar que la serie puede ser estudiada con los modelos Arch o Garch. Algunos efectos externos sobre la serie  $Z_t$  podrían ser los siguientes:

Vamos a analizar algunos años de la serie de tiempo precio externo del café colombiano:

En los años 1980-1986: En el mes de octubre se firma un gran acuerdo entre productores y consumidores de café. En ese tiempo los precios oscilaron dentro

de franjas pactadas hasta febrero de 1986, mes en que la sequía de Brasil ocasionó una nueva disparada de los precios.

En el año 2002: los datos históricos comienzan desde 02 de enero del 2002. Para esta fecha la oferta de café mostraba cinco años consecutivos superando ampliamente el consumo, a pesar de darse en una época de crisis mundial ; para muchos analistas , el año considerado como el más devastador en las historia del grano. Asimismo la producción total fue suficiente para atender la exportación y el consumo interno.

En el año 2003: comienza con una volatilidad elevada proveniente del año anterior, pero en descenso hasta el mes de marzo donde empieza a mantener una volatilidad más constante comparada con la de años anteriores: teniendo en cuenta que esta presenta elevaciones cuando el precio del café está bajando y viceversa.

En el año 2004 : comienza con un alza en los precios del café, mostrando volatilidades más bruscas, con descenso mucho más rápido comparado con el año anterior en el que la oferta estaba por debajo de lo que había venido presentando últimamente como efecto de la reducción de la producción de Brasil.

En el año 2005: vemos un alza en los precios del café para los primeros meses, sin embargo para mediados del mes de marzo comienza un descenso, hasta el mes de julio.

En el año 2006: comienza con un descenso después del buen precio alcanzado en el año 2005. En enero siguieron sólidos los precios del café donde se intensificaron los cambios en la volatilidad.

En al año 2007: sigue un aumento en el precio externo del café. En los años 2008-2009-2010-2011-2012 el precio del café sigue aumentando. En los años 2013-2014 el precio el café baja bruscamente, mostrando mucha volatilidad. En los años 2014- 2015 el precio externo del café tuvo una leve mejora.

Algunos de los factores externos o de la economía mundial que pudieron afectar estas volatilidades fueron:

La volatilidad de los precios apartir de 1990 pone de relieve un importante cambio en el sector cafetero mundial. Por una parte, la demora en la respuesta de los precios a impactos exógenos tales como choques climáticos se hizo mucho más corta y, por otra parte, por muy fuerte que sea la reacción, no persiste durante mucho tiempo. Los factores responsables de la excesiva volatilidad de los precios

del café deberían estudiarse inicialmente en términos de factores fundamentales del mercado, en especial relacionados con la oferta. La oferta está con frecuencia influenciada por factores exógenos relacionados con el clima en cuanto que un período de oferta escasa podrá seguirle un período de exceso de producción y viceversa. Los cambios en los factores fundamentales del mercado, podrán, por tanto favorecer o impedir que surjan factores especulativos.

La sobreoferta mundial desde los años 2002 en adelante del grano que ha estado explicada principalmente por el aumento de la producción en Brasil y Vietnam, comenzó a mostrar un cambio en tendencia.

Antes de 1989 la principal fuente de volatilidad del precio era la ocurrencia de fenómenos climáticos en el Brasil y que se conocía como heladas. En la medida que ese país fue trasladando su caficultura hacia zonas menos susceptibles de este fenómeno entonces aparecieron nuevas fuentes de volatilidad no solo climáticas sino de una índole más económica, como la mayor presencia de especuladores en los mercados del producto. Al mismo tiempo el clima se volvió aún más impredecible que en el pasado y aparecieron nuevos productores de bajos costos que han llenado de más incertidumbre al comportamiento de los precios de este producto en los mercados mundiales.

Por otra parte, han ocurrido cambios muy importantes por el lado de la demanda de café. Sin duda este producto a pesar de ser un commodity ha logrado un grado de diferenciación muy importante y mucho mayor que en otros productos de igual naturaleza. La mayor diferenciación ha llevado a una gran segmentación del mercado del café a nivel mundial.

Resumiendo, estos son algunos factores que han afectado la volatilidad del precio del café colombiano:

- La expansión cafetera a nivel de la producción, especialmente en Vietnam
- El aumento importante en la siembra de árboles por parte de la industria brasilera del café.
- El cambio tecnológico que ha permitido mejorar el control de plagas y modificar la participación de los cafés suaves.
- Los movimientos en las tasas de cambio y en especial la reevaluación del peso colombiano.
- El cambio climático, períodos de mucha sequía seguidos de períodos de muchas lluvias.
- Crecimiento lento del consumo.

4.2.1 Correlograma de retornos  $\gamma_t$  (a la serie precio externo del café colombiano se le aplicó un logaritmo y una diferencia).

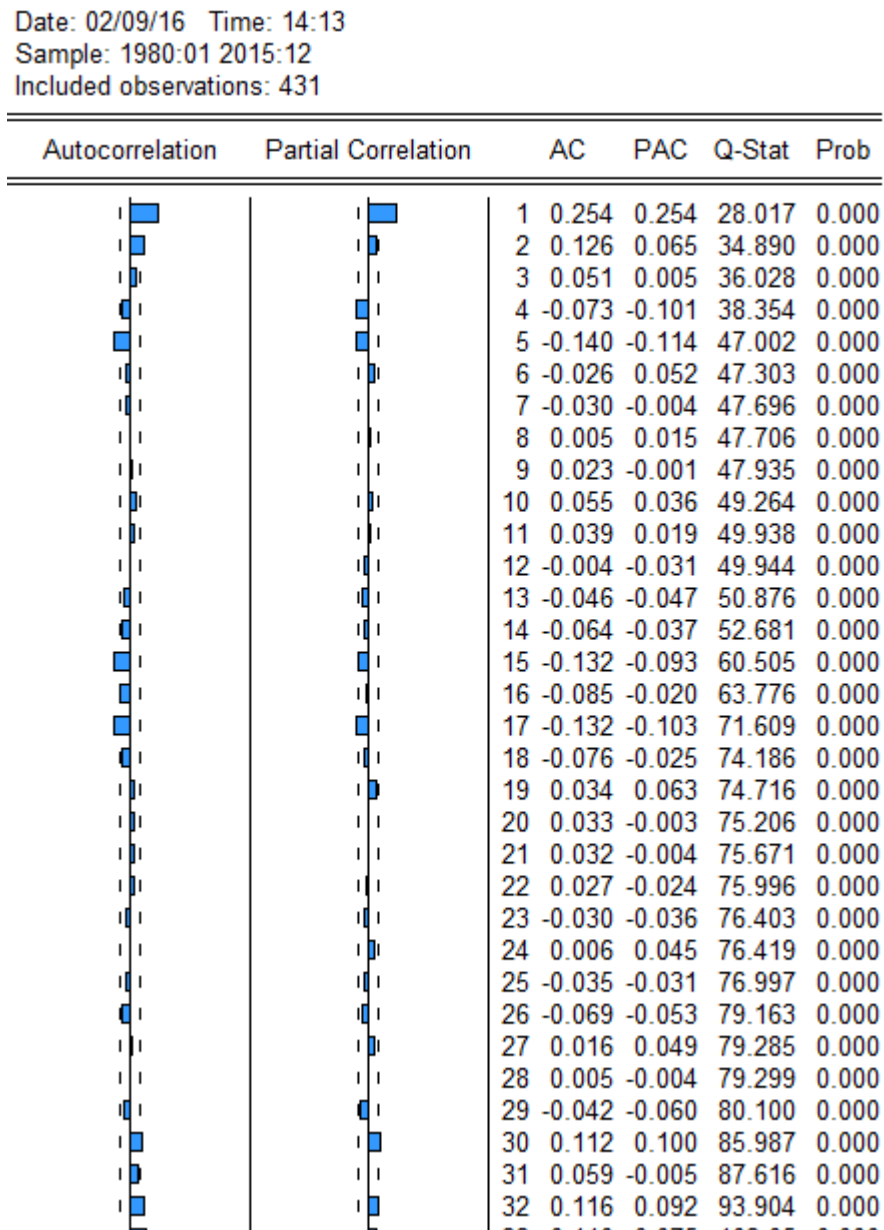


Figura 4.2.1 Correlograma de la serie de los retornos  $\gamma_t$ .

Observamos que tanto la FAC como la FAP caen rápidamente a cero, sospechando que la serie es estacionaria.

ADF Test Statistic	-10.11160	1% Critical Value*	-3.4478
		5% Critical Value	-2.8686
		10% Critical Value	-2.5705

\*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(X)  
 Method: Least Squares  
 Date: 02/09/16 Time: 14:17  
 Sample(adjusted): 1980:07 2015:12  
 Included observations: 426 after adjusting endpoints

Tabla 4.3 : Test de Dickey Fuller aplicado a la serie de los retornos

El test de Dickey Fuller aplicado a la serie de los retornos verifica la no existencia de raíces unitarias.

Como el estadístico en valor absoluto 10.11 es mayor que los valores críticos al 1%, 5%, 10% concluimos que la serie precio externo del café colombiano después de haber aplicado logaritmo y una diferencia es estacionaria.

#### 4.2.2 histograma de la serie DLOG (Z)

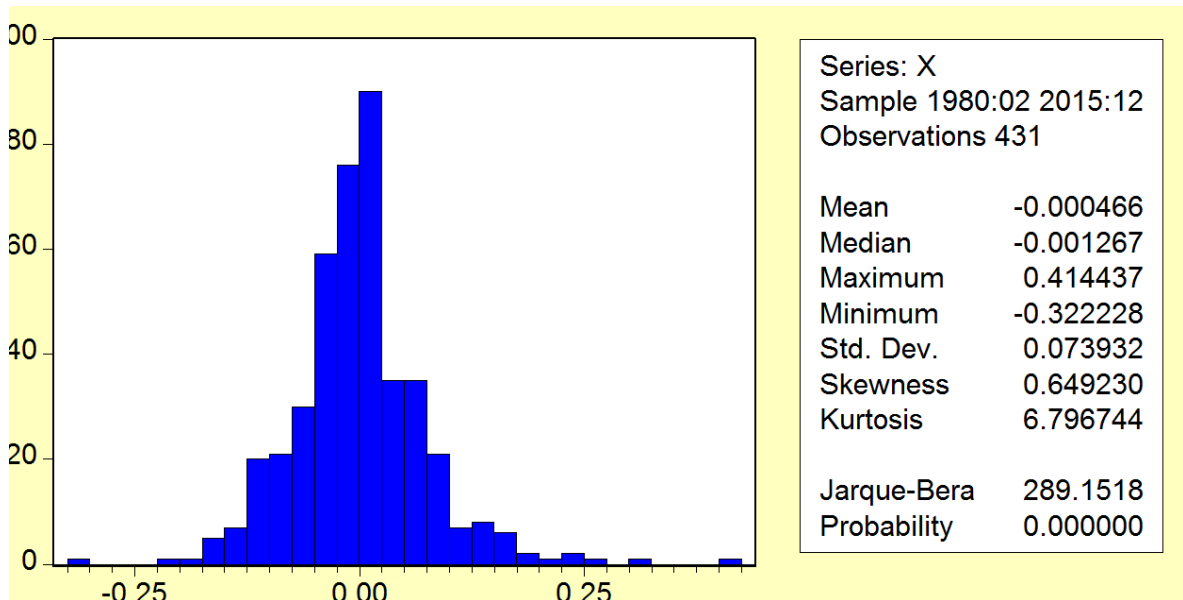


Figura 4.2.2 histograma de la serie DLOG (Z)

Según la figura 4.3.3 y el resumen estadístico, se Observa que la serie presenta elevada kurtosis ( $K \cong 6,8 > 3$ ) También que las colas son muy pequeñas.

Recordemos que un valor grande de la kurtosis significa que los datos están muy concentrados cerca de la media de la distribución coexistiendo al mismo tiempo con una relativa elevada frecuencia de datos alejados de la misma. Esto explica una forma de distribución de frecuencias con colas muy pesadas y con un centro muy elevado.

#### 4.4 Prueba de Heterocedasticidad de la serie $Z_t$ : precio externo del café colombiano.

Dependent Variable: Z  
Method: Least Squares  
Date: 02/09/16 Time: 16:18  
Sample: 1980:01 2015:12  
Included observations: 432

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	137.4706	2.463661	55.79929	0.0000
R-squared	0.000000	Mean dependent var	137.4706	
Adjusted R-squared	0.000000	S.D. dependent var	51.20624	
S.E. of regression	51.20624	Akaike info criterion	10.71191	
Sum squared resid	1130116.	Schwarz criterion	10.72133	
Log likelihood	-2312.773	Durbin-Watson stat	0.049612	

Tabla 4.4: Regresión de la serie contra la constante

Lo primero que hacemos es una regresión de la serie: precio externo del café colombiano contra una constante.

Obteniendo que la constante es estadísticamente significativa y con un Durbin Watson prácticamente cero.

#### 4.3 Test de Heterocedasticidad para la serie precio externo del café colombiano

ARCH Test:

F-statistic	3476.800	Probability	0.000000
Obs*R-squared	383.6604	Probability	0.000000

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 02/09/16 Time: 16:31

Sample(adjusted): 1980:02 2015:12

Included observations: 431 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	144.9206	79.97027	1.812182	0.0707
RESID^2(-1)	0.943759	0.016006	58.96440	0.0000

Tabla 4.5: Prueba de Heterocedasticidad para la serie precio externo  
Del café colombiano.

$H_0$  : No hay efectos Arch

$H_1$  : Si existen efectos Arch

No hay evidencia estadística para rechazar la hipótesis nula, luego si existen efectos Arch. Esto nos conduce a pensar que la serie precio externo del café colombiano puede ser modelada por modelos Arch, Garch, Igarch, Tarch.

Los parámetros cumplen la condición, la constante es mayor que cero y el parámetro sujeto a la regresión es inferior a uno.

## 5. MODELACIÓN DE LA SERIE PRECIO EXTERNO DEL CAFÉ COLOMBIANO USANDO EVIEWS 3.1

### 5.1 Ajuste del Modelo AR 1 + ARCH (1)

Dependent Variable: DLOG(Z)  
 Method: ML - ARCH  
 Date: 02/06/16 Time: 08:11  
 Sample(adjusted): 1980:03 2015:12  
 Included observations: 430 after adjusting endpoints  
 Convergence achieved after 11 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.184570	0.043652	4.228251	0.0000
Variance Equation				
C	0.003927	0.000260	15.12818	0.0000
ARCH(1)	0.252281	0.054063	4.666462	0.0000
R-squared	0.059719	Mean dependent var	-0.000414	
Adjusted R-squared	0.055315	S.D. dependent var	0.074011	
S.E. of regression	0.071935	Akaike info criterion	-2.468453	
Sum squared resid	2.209546	Schwarz criterion	-2.440101	
Log likelihood	533.7174	Durbin-Watson stat	1.874535	
Inverted AR Roots	.18			

Tabla 5.1 Ajuste del Modelo AR ( 1 ) - ARCH ( 1 ) a la serie de los retornos.

Destacamos en este modelo que tanto la constante como el parámetro son estadísticamente significativos. Además el criterio Akaike es de -2.468453 y el Durbin Watson es de 1.874535

El proceso es estacionario.

$$\gamma_t = 0.1845 \gamma_{t-1} + a_t$$

$$a_t = \sqrt{h_t} \varepsilon_t$$

$$h_t = 0.0039 + 0.2529 a_{t-1}^2$$



### 5.1.1 Ajuste del Modelo AR + GARCH

Dependent Variable: DLOG(Z)				
Method: ML - ARCH				
Date: 02/13/16 Time: 16:02				
Sample(adjusted): 1980:03 2015:12				
Included observations: 430 after adjusting endpoints				
Convergence achieved after 20 iterations				
	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.223676	0.060285	3.710328	0.0002
Variance Equation				
C	0.001138	0.000408	2.788375	0.0053
ARCH(1)	0.172258	0.041542	4.146640	0.0000
GARCH(1)	0.603474	0.114786	5.257375	0.0000
R-squared	0.063630	Mean dependent var	-0.000414	
Adjusted R-squared	0.057035	S.D. dependent var	0.074011	
S.E. of regression	0.071869	Akaike info criterion	-2.505515	
Sum squared resid	2.200357	Schwarz criterion	-2.467712	
Log likelihood	542.6857	Durbin-Watson stat	1.959259	
Inverted AR Roots	.22			

Tabla 5.1.1: Ajuste del modelo AR (1) - GARCH (1,1) a la serie de los retornos.

Destacamos en este modelo que tanto los 3 parámetros como la constante son estadísticamente significativos. Además el criterio Akaike es de -2.505515 y el Durbin Watson es de 1.959259.

El proceso es estacionario.

El modelo es:

$$\gamma_t = 0.2237\gamma_{t-1} + a_t$$

$$a_t = \sqrt{h_t} \varepsilon_t$$

$$h_t = 0.00114 + 0.1722a_{t-1}^2 + 0.6035h_{t-1}$$

La varianza incondicional está dada por:

$$\widehat{var}(\hat{a}_t) = \frac{0.00114}{1-0.1722-0.6035} = \frac{0.00114}{0.2243} = 0.0050824$$

### 5.1.2 Ajuste del Modelo AR + EGARCH

Dependent Variable: DLOG(Z)  
Method: ML - ARCH  
Date: 02/06/16 Time: 08:13  
Sample(adjusted): 1980:03 2015:12  
Included observations: 430 after adjusting endpoints  
Convergence achieved after 65 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.250638	0.060800	4.122338	0.0000
Variance Equation				
C	-1.632999	0.480502	-3.398527	0.0007
RES /SQR[GARCH](1)	0.329872	0.063984	5.155538	0.0000
RES/SQR[GARCH](1)	0.140042	0.037790	3.705765	0.0002
EGARCH(1)	0.742478	0.084123	8.826092	0.0000
R-squared	0.064544	Mean dependent var	-0.000414	
Adjusted R-squared	0.055740	S.D. dependent var	0.074011	
S.E. of regression	0.071918	Akaike info criterion	-2.514429	
Sum squared resid	2.198208	Schwarz criterion	-2.467175	
Log likelihood	545.6022	Durbin-Watson stat	2.017082	
Inverted AR Roots	.25			

Tabla 5.1.2 Ajuste del modelo AR (1) + EGARCH (1,1) a la serie de los retornos

Destacamos en este modelo que tanto los 3 parámetros como la constante son estadísticamente significativos. Además el criterio Akaike es de -2.514429 y el Durbin Watson es de 2.017082.

El modelo es:

$$\gamma_t = 0.25\gamma_{t-1} + a_t$$

$$a_t = \sqrt{h_t}\varepsilon_t$$

$$\ln h_t = -1.63 + 0.74\ln h_{t-1} + 0.33|\varepsilon_{t-1}| + 0.142$$

### 5.1.3 Ajuste del Modelo AR + TARCH

Dependent Variable: DLOG(Z)  
Method: ML - ARCH  
Date: 02/06/16 Time: 08:12  
Sample(adjusted): 1980:03 2015:12  
Included observations: 430 after adjusting endpoints  
Convergence achieved after 15 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.232869	0.061095	3.811564	0.0001
Variance Equation				
C	0.001736	0.000529	3.279313	0.0010
ARCH(1)	0.297671	0.098602	3.018903	0.0025
(RESID<0)*ARCH(1)	-0.240267	0.095648	-2.512008	0.0120
GARCH(1)	0.474217	0.137380	3.451872	0.0006
R-squared	0.064105	Mean dependent var	-0.000414	
Adjusted R-squared	0.055296	S.D. dependent var	0.074011	
S.E. of regression	0.071935	Akaike info criterion	-2.514785	
Sum squared resid	2.199240	Schwarz criterion	-2.467532	
Log likelihood	545.6789	Durbin-Watson stat	1.979046	
Inverted AR Roots	.23			

Tabla 5.1.3 Ajuste del Modelo AR ( 1 ) + TARCH (1,1) a la serie de los retornos

$$\gamma_t = 0.233\gamma_{t-1} + a_t$$

$$a_t = \sqrt{h_t} \varepsilon_t$$

$$h_t = 0.0017 + 0.298a_{t-1}^2 - 0.24a_{t-1}^2d_{t-1} + 0.474h_{t-1}$$

Destacamos en este modelo que tanto los 3 parámetros como la constante son estadísticamente significativos. Además el criterio Akaike es de -2.514785 y el Durbin Watson es de 1.979046.

### 5.1.4 Selección del Modelo

MODELO	AKAIKE
AR(1) +ARCH 1	-2.4684
AR(1) +GARCH (1,1)	-2.5055
AR(1) +TARCH(1,1)	-2.5147
AR(1) +EGARCH(1,1)	-2.5144

Tabla 5.1.4: Elección del mejor modelo

El mejor modelo es el asimétrico AR (1) + TARCH (1,1) por tener el menor criterio Akaike.

## 5.2 Pronósticos dados por el modelo TARCH (1,1)

PRONÓSTICOS EViews 3.1 MODELO TARCH (1,1) SERIE: PRECIO EXTERNO DEL CAFÉ COLOMBIANO

SERIE ORIGINAL	PRONÓSTICO
Z	ZF
170.67	
166.83	
186.38	165.95
182.81	191.25
125.77	123.04
132.02	125.90
162.37	133.52
211.18	170.39
216.11	224.51
217.53	217.27
195.82	217.86
193.71	191.08
210.49	193.22
194.27	214.60
201.83	190.68
206.61	203.63
191.8	207.74
185.82	188.51
177.11	184.45
156.88	175.14
156.43	152.51
152.52	156.33
150.95	151.62
145.95	150.59
147.4	144.81
136.6	14.77
142.69	134.20
138.3	144.15
139.63	137.30
PREDICCIÓN ENERO DEL 2016	139.94

Tabla 5.2: Pronósticos del modelo TARCH (1,1) para el mes de Enero del 2016

## 5.3 comparativo Pronósticos

Porcentaje de Error del Modelo TARCH (1,1) y de los arrojados por el aplicativo en Excel para el mes de Enero del 2016, serie de tiempo: precio externo del café colombiano.

	% de Error	Valor real para Enero 2016. Serie: precio externo del café colombiano
Modelo Tarch(1,1) Eviews 3.1	2.6%	136.30 centavos dólar
Suavización Exponencial	2.38%	
Media Móviles	2.86%	
Método de regresión lineal simple	15.49%	

Tabla: 5.3 comparativo pronósticos

## 5.4 Modelación en Matlab

### 5.4.1 Modelo Arch1

Did you mean:

```
>> estm=estimate mdl,X)
```

```
GARCH(0,1) Conditional Variance Model:
```

```
-----  
Conditional Probability Distribution: Gaussian
```

Parameter	Value	Standard Error	t Statistic
Constant	0.00392304	0.000246235	15.9321
ARCH{1}	0.284702	0.0463398	6.14379

```
estm =
```

```
GARCH(0,1) Conditional Variance Model:
```

```
-----  
Distribution: Name = 'Gaussian'
```

```
P: 0
```

```
Q: 1
```

```
Constant: 0.00392304
```

```
GARCH: {}
```

```
ARCH: {0.284702} at Lags [1]
```

Figura 5.4.1 parámetros modelo Arch 1 Matlab

#### 5.4.1.1 Criterio Akaike

Modelo :	CRITERIO AKAIKE
ARCH 1	3298 .87

Tabla 5.4.1. Criterio akaike modelo Arch 1

#### 5.4.2 Modelo Garch (1,1)

```

GARCH(1,1) Conditional Variance Model:
-----
Conditional Probability Distribution: Gaussian

Parameter      Value      Standard      t
              Value      Error      Statistic
-----
Constant      0.00151716    0.00041625    3.64482
GARCH{1}      0.496221     0.111485     4.45101
ARCH{1}       0.224237     0.0428301    5.2355

estm =

GARCH(1,1) Conditional Variance Model:
-----
Distribution: Name = 'Gaussian'
          P: 1
          Q: 1
Constant: 0.00151716
GARCH: {0.496221} at Lags [1]
ARCH: {0.224237} at Lags [1]

```

Figura: 5.4.2 Parámetros modelo Garch (1,1) Matlab

#### 5.4.2.1 Criterio Akaike

Modelo :	CRITERIO AKAIKE
GARCH (1,1)	3297.26

Tabla 5.4.2 Criterio akaike modelo Garch (1,1)

### 5.4.3 Modelo Tarch (1,1)

```
GJR(1,1) Conditional Variance Model:
-----
Conditional Probability Distribution: Gaussian

Parameter      Value      Standard      t
      Error      Statistic
-----
Constant      0.00188514    0.00047782     3.94528
GARCH{1}      0.422089      0.118267     3.56896
ARCH{1}       0.330673      0.0880232     3.75666
Leverage{1}   -0.213383      0.0869055    -2.45534

stm =

GJR(1,1) Conditional Variance Model:
-----
Distribution: Name = 'Gaussian'
      P: 1
      Q: 1
Constant: 0.00188514
GARCH: {0.422089} at Lags [1]
ARCH: {0.330673} at Lags [1]
Leverage: {-0.213383} at Lags [1]
```

Figura: 5.4.3 Parámetros modelo Tarch (1,1) Matlab

#### 5.4.3.1 Criterio Akaike

Modelo :	CRITERIO AKAIKE
TARCH (1,1)	3297.26904

Tabla 5.4.3: criterio akaike modelo TARCH (1,1)

#### 5.4.4 Modelo Egarch (1,1)

```

EGARCH(1,1) Conditional Variance Model:
-----
Conditional Probability Distribution: Gaussian

Parameter      Value      Standard      t
              Value      Error      Statistic
-----
Constant      -1.62535    0.439728    -3.69626
GARCH{1}      0.690183    0.0809605    8.52493
ARCH{1}       0.422786    0.0629893    6.71204
Leverage{1}   0.129044    0.0390251    3.3067

estm =

EGARCH(1,1) Conditional Variance Model:
-----
Distribution: Name = 'Gaussian'
          P: 1
          Q: 1
Constant: -1.62535
      GARCH: {0.690183} at Lags [1]
      ARCH: {0.422786} at Lags [1]
Leverage: {0.129044} at Lags [1]

```

Figura 5.4.4 Parámetros modelo Egarch ( 1,1)

##### 5.4.4.1 Criterio Akaike

Modelo :	CRITERIO AKAIKE
EGARCH (1,1)	3297.99

Tabla 5.4.4: Criterio akaike Modelo EGARCH (1,1)

##### 5.4.5 Selección del mejor Modelo

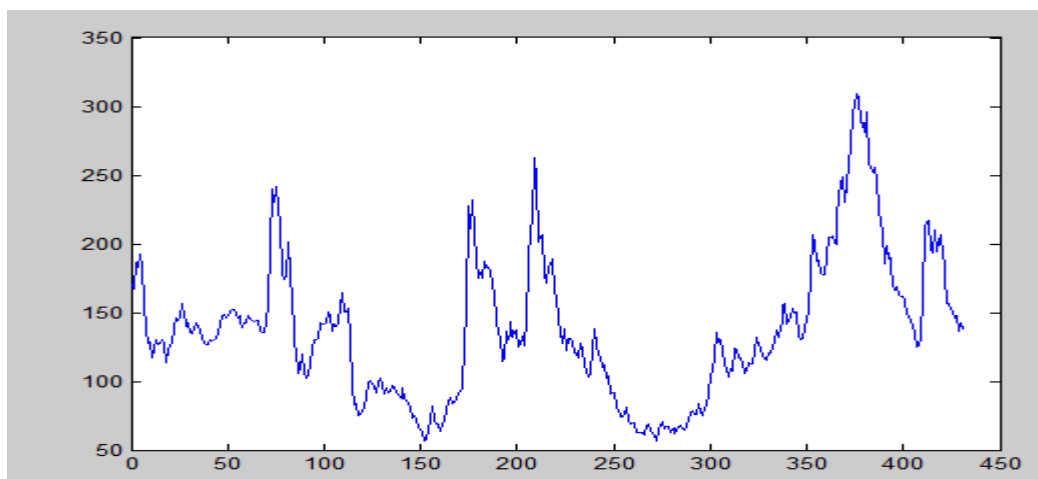
Modelo	Akaike
Arch 1	3298.8712
Garch (1,1)	3297.26904
Tarch (1,1)	3297.2655
Egarch (1,1)	3297.99

Tabla 5.4.5 resumen de los criterios akaike para modelos

Seleccionamos el modelo TARCH (1,1) por tener el valor del criterio Akaike más pequeño.



5.4.5 Gráfica de la serie precio externo del café colombiano vs precio en centavos de dólar en Matlab.



Gráfica 5.4.5: gráfica del precio externo del café colombiano

5.4.5 Gráfica de la serie después de aplica logaritmos y una diferencia en Matlab.

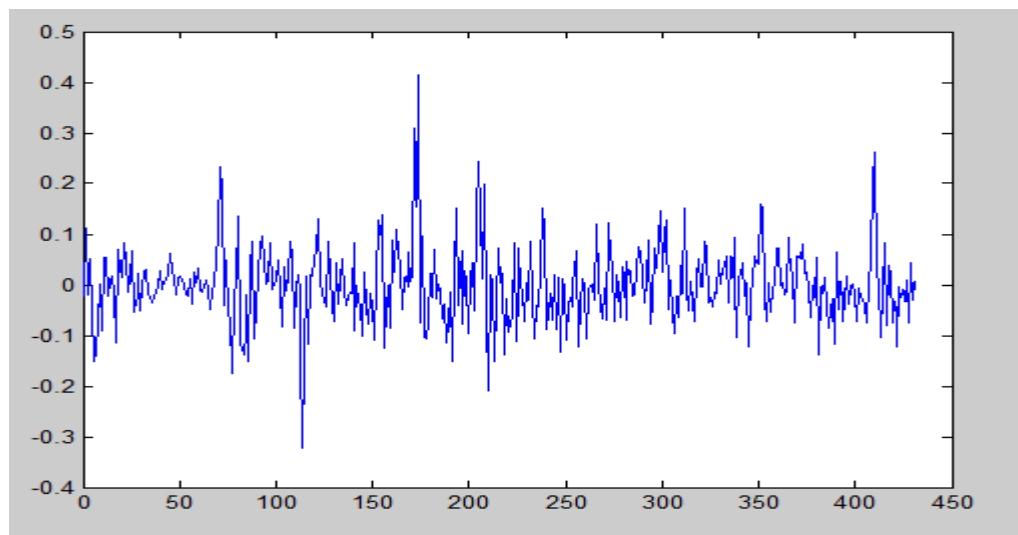


Figura 5.4.6. gráficas de la serie después de haber aplicado logaritmo y una diferencia.

```
Did you mean:  
>> adftest(X)  
  
ans =  
  
1
```

Figura: 5.4.7 Test de Dickey Fuller para la serie precio externo del café después de aplicar logaritmos y una diferencia.

Según este test de Dickey Fuller la serie precio externo del café colombiano es estacionario después de haber aplicado logaritmos y una diferencia. Matlab

Test de efectos Arch para la serie precio externo del café colombiano después de haber aplicado logaritmos y una diferencia.

```
>> archtest(Z)  
  
ans =  
  
1
```

Figura:5.4.8 test de efectos arch de la serie precio externo del café colombiano. Matlab

Según este test de la serie Precio externo del café colombiano presenta efectos Arch, es decir de Heterocedasticidad condicionada.

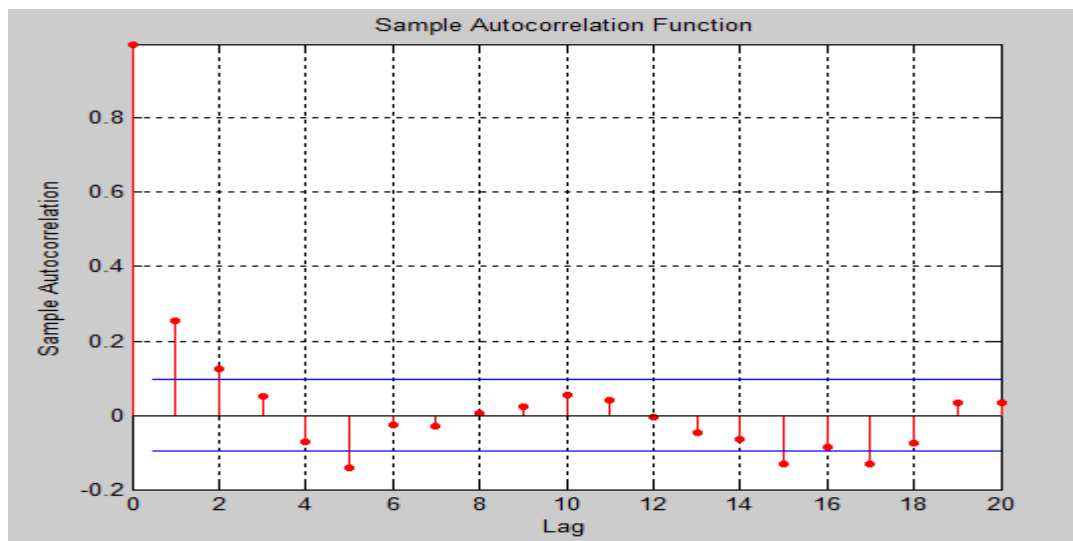
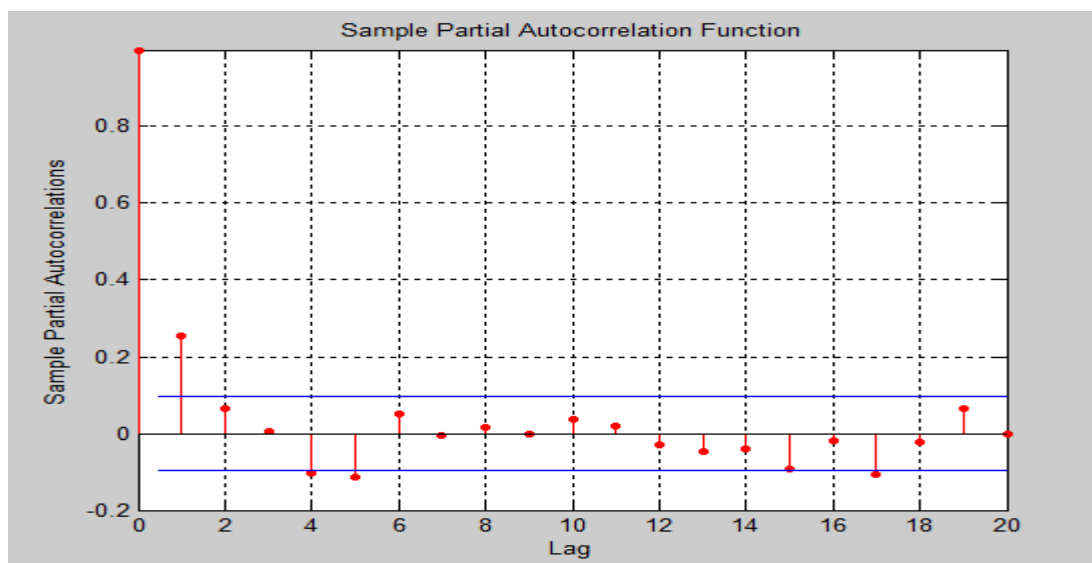


Figura 5.4.9: Gráfica de Auto correlación FAC de la serie precio externo del café colombiano después de haber aplicado logaritmo y una diferencia.



Gráfica: 5.5 autocorrelación parcial de la serie precio externo del café Colombiano.

5.6 FAC y FACP de los residuos de la serie precio externo del café colombiano modelo Tarch (1,1) Matlab

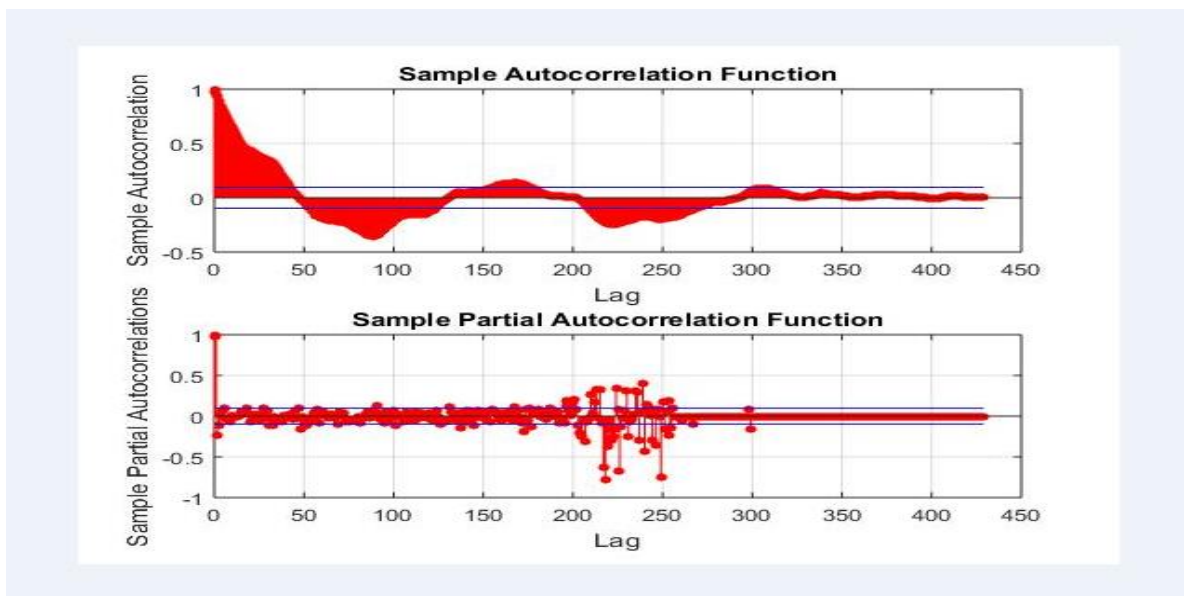


Figura 5.6: FAC Y FACP de residuos modelo tarch (1,1)

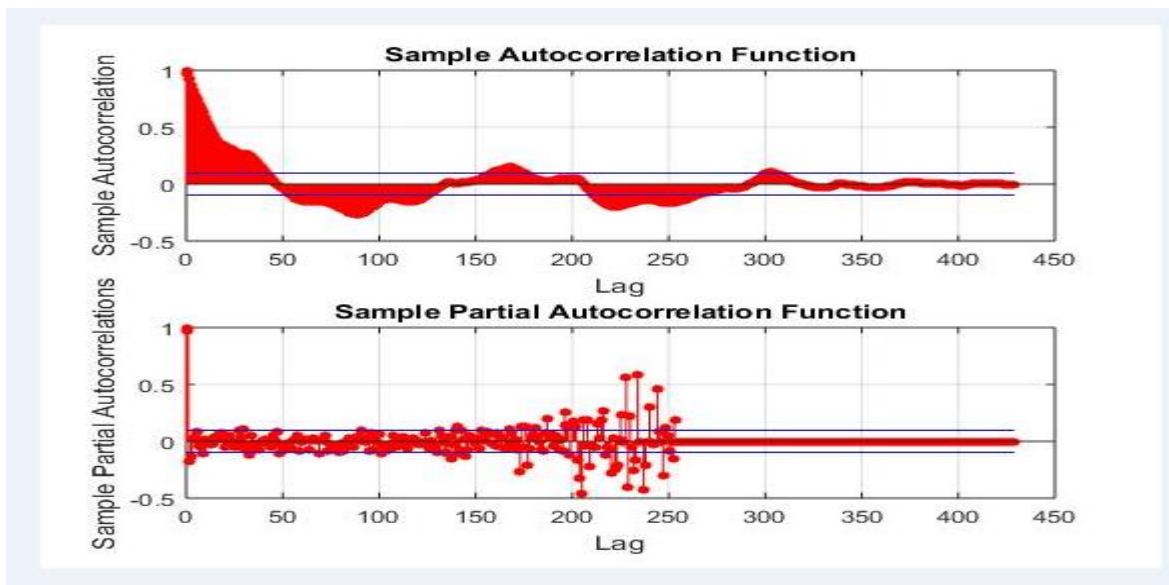


Figura 5.7 :Función de autocorrelación simple y parcial del modelo Tarch (1,1)

#### 5.4.6.8. Resumen de criterios del Modelo Tarch (1,1)

Criterios	
Residuos	1
Residuos al cuadrado	1
Jarque-Bera	1
Akaike	3328.36738

Dickey Fuller	1
---------------	---

Tabla 5.4.6.8: criterios del modelo TARCH (1,1)

#### 5.4.6.9. Pronósticos Modelo Tarch (1,1)

Mes	Pronóstico
Enero 2016	139.560015
Febrero	141.919076
Marzo	144.278138
Abril	146.637199
Mayo	148.996259
Junio	151.355319

Tabla 5.4.6.9: Pronósticos Modelo Tarch (1,1)

#### 5.4.7 % de Error relativa del modelo Tarch (1,1) usando Matlab

	% de Error	Valor real para Enero 2016. Serie: precio externo del café colombiano
Modelo Tarch(1,1) Matlab	2.39%	136.30 centavos dólar

Tabla: 5.4.7 % de Error relativa del modelo Tarch (1,1) usando Matlab

# CONCLUSIONES

- La serie de tiempo precio externo del café colombiano muestra presencia de Efectos Arch, esto nos indica que la volatilidad de la serie no es constante a lo largo del tiempo.
- La serie Precio externo del café colombiano es asimétrica, leptocúrtica (una de las principales características de las series financieras). El test de Jarque Bera nos muestra que la serie no sigue una Distribución Normal.
- La serie de tiempo precio externo del café colombiano puede ser modelada por un modelo AR (1) +Tarch (1,1), que refleja muy bien el comportamiento de la serie. Puede servir de herramienta para la predicción de la volatilidad con el objeto de utilizarla en negociaciones del café.
- El porcentaje de error relativo del Modelo Tarch (1,1) modelado con Eviews 3.1 fue de 2.6% para el mes de Enero del 2016.
- El método de Suavización Exponencial (con un  $\alpha = 0.7$ ), del aplicativo en Excel, fue el más aproximado en el pronóstico para el mes de enero del 2016, pues arrojó un porcentaje de error relativo del 2.38%, frente a un 2.86% del método de Medias Móviles y a un 15.49% del método de Regresión lineal simple.
- El porcentaje de error relativo del Modelo Tarch (1,1) modelado con Matlab fue del 2.39%.
- El aplicativo en Excel nos arrojó un buen pronóstico para enero del 2016, destacando los métodos de medias móviles y de suavización exponencial además de ser un buen recurso didáctico para la enseñanza de series de tiempo.

## BIBLIOGRAFÍA:

- Pérez Ramírez Fredy Ocaris, 2006, Modelación de la volatilidad y pronóstico del precio del café. Revista de ingenierías Universidad de Medellín.
- Peña Daniel  
Series de Tiempo, Capítulo 15: Análisis Univariante de series temporales (página 549 – 585). Alianza Editorial.
- Urbisaia L. Heriberto, Brufman Z. Juana  
Análisis de Series de Tiempo: Univariadas y Multivariadas. Segunda edición Ediciones Cooperativas.
- Quesada Pegalajar Manuel, Análisis de series. Modelos Heterocedásticos Universidad Granatensis, España. Trabajo fin de master.
- García Pérez Julián, San Román Antonio Pulido, guía para la elaboración de modelos econométricos con Eviews. Modelos Econométricos.
- Gujarati N Damodar, Porter C. Dawn, quinta edición, Mc Graw Hill, 2009, Capítulo 21: econometría de series de tiempo: algunos conceptos básicos, Capítulo 22: econometría de las series de tiempo: pronóstico.
- Morettin A. Pedro, Toloí C. M Clélia, segunda edición, 2006, Capítulo 14: Modelos no lineales, páginas 361-410 .Análisis de series Temporales. ABE-Projeto Fisher.
- Contreras Espinoza Adriana Elisa, El caos y la caracterización de series de tiempo a través de técnicas de la dinámica no lineal. 2004. Universidad Nacional Autónoma de México. Campus Aragón.

## ANEXO

### Manual de Usuario Aplicativo en Excel

#### Descripción del Aplicativo

#### Tutorial: Aplicativo Series de Tiempo

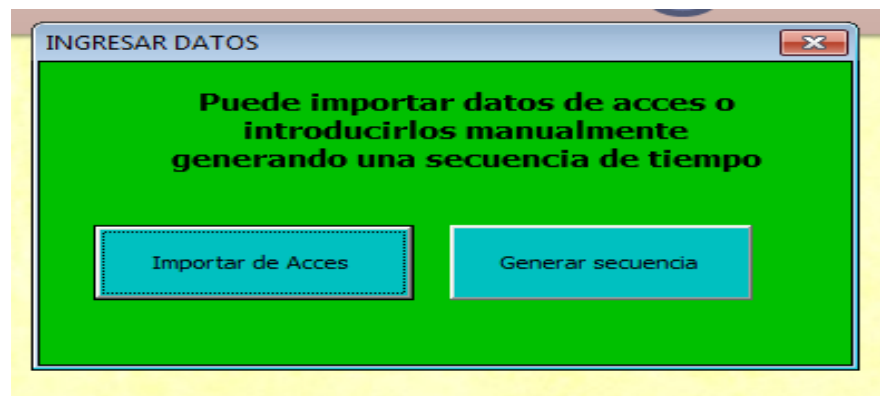
#### Menú principal



#### Elementos del menú principal:

1. Datos 2. Modelo de regresión lineal simple 3. Media móviles 4. Suavización Exponencial 5. Graficar 6. Pronósticos 7. Limpiar pantalla 8. Inicio

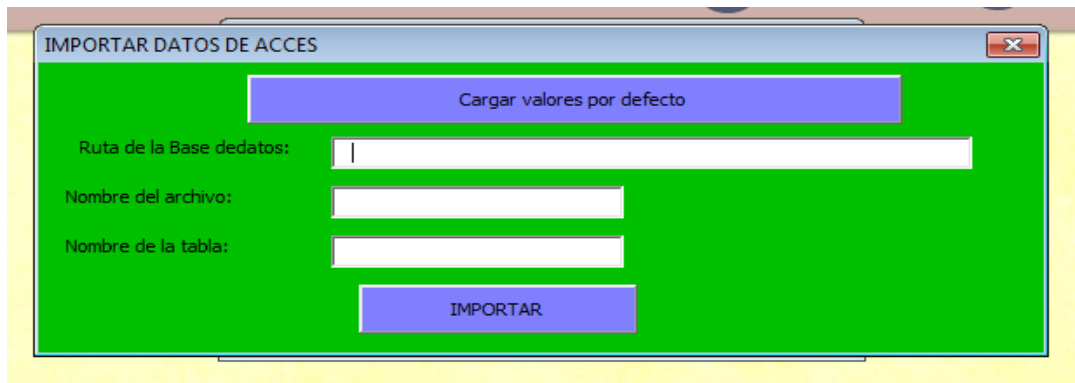
1. Mediante el botón Datos se cargan los datos de la siguiente forma:



Estos datos se pueden importar de una base de datos en Acces, de lo contrario los importamos con el botón generar secuencia.

Si los importamos desde Acces nos aparece:





Al darle cargar valores por defecto e importar podemos automáticamente cargar la base de datos de Acces.

Si le damos generar secuencia, el aplicativo nos pregunta la fecha de inicio y la fecha de fin. Luego podemos copiar y pegar datos de Excel.

## 2. Modelo de Regresión lineal simple:

Estudia la relación lineal entre la variable respuesta ( Y ) y la variable represora ( X ) , a partir de la muestra  $\{(x_i, Y_i)\}_{i=1}^n$  , que sigue el siguiente modelo:

$$Y_i = \alpha_0 + \alpha x_i + \varepsilon_i \quad i=1, 2, 3 \dots n$$

DATOS		
?		
AÑO	VALOR	MIN CU
1913	15	-12
1914	15	-11
1915	14	-9
1916	14	-7
1917	12	-6
1918	16	-4
1919	27	-2
1920	21	-1
1921	15	1

El signo ? nos informa la teoría del método de regresión lineal simple

## 3. Medias Móviles

Este método calcula las medias móviles. El signo ? nos informa la teoría del método de medias móviles. Cuando presiono el botón aparecen los datos del ajuste por medias móviles.

MEDIA MOVIL	
	15
	14
	13
	14
	10
	21
	21

#### 4. Suavización Exponencial

Primero aparece un cuadro con la pregunta: cuál es la constante de suavización que vamos a utilizar?

Suavización Exponencial

CONSTANTE DE SUAVIZACIÓN

0.6

*La constante de suavización debe ser un valor entre cero y uno. Separe las cifras decimales con una coma*

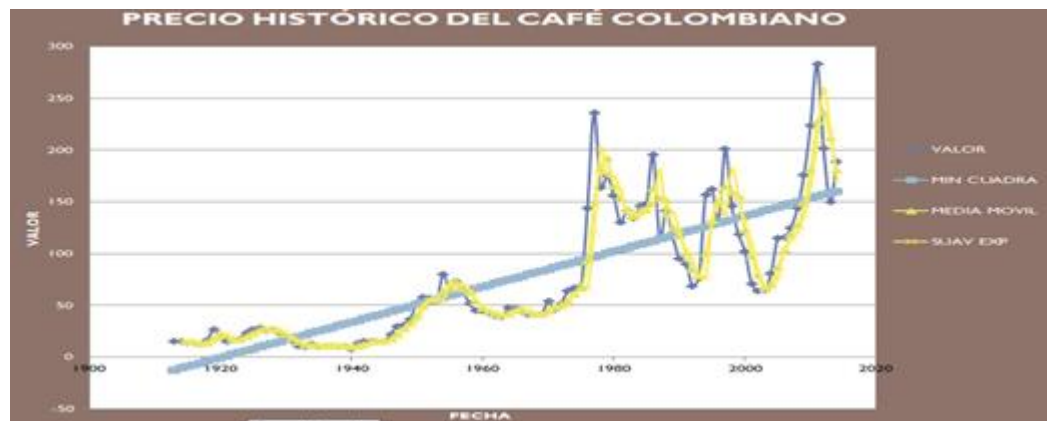
ACEPTAR

Esta constante está entre 0 y 1. El valor se debe colocar con coma, no con punto.

SUAV EX	
	15
	15
	14
	14
	13
	15
	23
	22
	17

#### 5. Graficar

Aparece una gráfica con el ajuste de mínimos cuadrados. También se visualiza la ecuación de la gráfica de mínimos cuadrados. Destacándose la pendiente y el intercepto con el eje y.



## 6. Pronósticos

Nos aparece un cuadro para seleccionar los pronósticos, que puede ser por: Medias Móviles, Regresión lineal simple, Suavización exponencial. Si seleccionamos el de medias móviles, entonces nos aparece el siguiente cuadro:

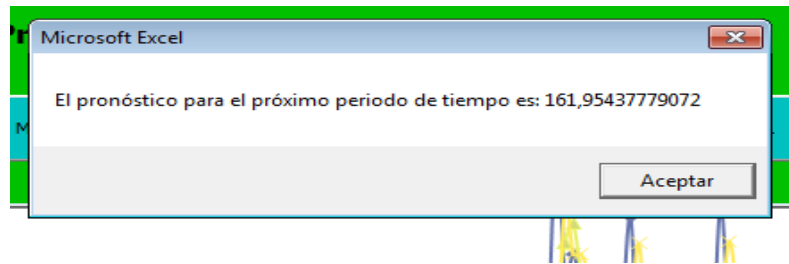
FECHA	MEDIA CALCULA	PESO

Al cargar valores, el aplicativo toma los tres últimos años antes del valor a pronosticar, pregunta los pesos, por si queremos darle más importancia a un dato en particular, luego nos arroja el pronóstico.

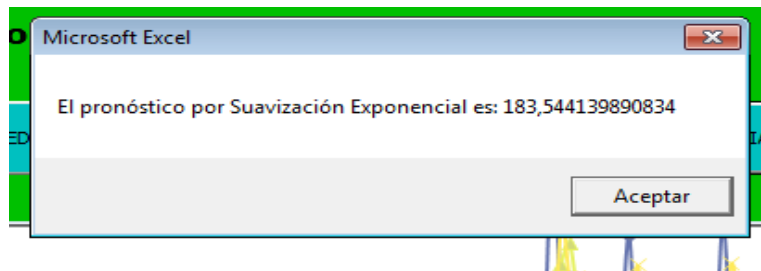
FECHA	MEDIA CALCULA	PESO
2012	236,33333333	1
2013	211,66666666	2
2014	180,33333333	2

**Valor para el próximo periodo:  
204,066666666667**

Al pronosticar por el método de regresión lineal simple, entonces aparece el valor del pronóstico en un recuadro:

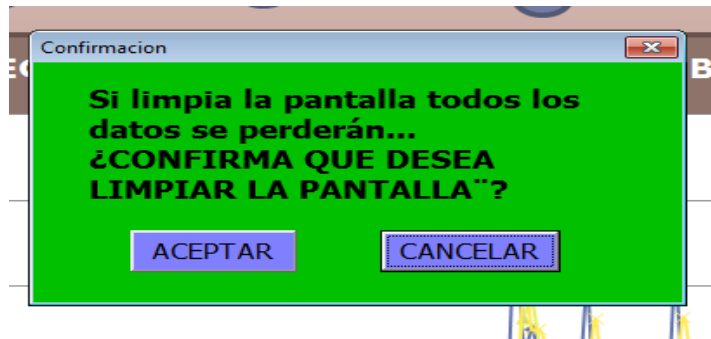


Al pronosticar con el método de suavización exponencial, nos aparece un recuadro:



## 7. Limpiar pantalla

Con esta tecla sale el siguiente cuadro:



Al aceptar, se borran automáticamente todos los cálculos hechos anteriormente y podemos realizar el análisis de otra serie de tiempo.

## 8. Inicio

Al apretar el botón, nos aparece lo siguiente:



La primera opción es conocer cómo funciona el programa, en ésta nos informamos de todos los pasos y características teóricas que tiene el aplicativo.

Al seleccionar menú, nos traslada al menú principal y podemos empezar el análisis de una nueva serie de tiempo.

### Resumen de cómo funciona el Aplicativo en Excel

Necesitamos una base de datos en Acces, en caso contrario entramos los datos manualmente. Luego al aplicativo importa los datos. Después podemos recorrer cada uno de los métodos. El primero es el de mínimos cuadrados, luego el de medias móviles, Suavización Exponencial, luego podemos graficar la base de datos.

Al aplicar la tecla pronósticos, el aplicativo nos pregunta con cuál de los métodos lo vamos a realizar, si es con el de mínimos cuadrados el aplicativo realiza el pronóstico y nos arroja el respectivo error, si es con el de media móviles, el aplicativo realiza el pronóstico y nos arroja el respectivo error, si es con el de suavización exponencial , el aplicativo realiza el pronóstico y nos arroja el respectivo error..

En la parte derecha de la pantalla, aparecen los pronósticos respectivos y podemos comparar.

Con la tecla limpiar pantalla lo que hacemos es borrar y dejamos el aplicativo listo para una nueva serie de tiempo.

Con la tecla inicio lo que hacemos es regresar al menú principal, ya que el aplicativo tiene una parte estructurada d conceptos teóricos, importantes para entender series de tiempo.

